Fate of Transcription Elongation Complexes Stalled by DNA Damage and Elongation Inhibitors

by

Rachel Krasich

Department of Biochemistry
Duke University

Date:_____

Approved:

_____
Kenneth N. Kreuzer, Supervisor

_____
Arno L. Greenleaf

_____
Meta J. Kuehn

_____
Christopher Nicchitta

_____
Pei Zhou

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Biochemistry in the Graduate School
of Duke University

2014

Abstract

Fate of Transcription Elongation Complexes Stalled by DNA Damage and Elongation Inhibitors

by

Rachel Krasich

Department of Biochemistry
Duke University

Date:_____
Approved:

_____
Kenneth N. Kreuzer, Supervisor

_____
Arno L. Greenleaf

_____
Meta J. Kuehn

_____
Christopher Nicchitta

_____
Pei Zhou

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry in the Graduate School of Duke University

2014

Abstract

Transcription, the essential process by which cells translate genetic information stored in DNA into RNA, is a highly regulated and discontinuous process. Elongation is frequently blocked by DNA damage, pause sites, or intrinsic or external inhibitors. Due to the essential nature of transcription, the cell has numerous ways of dealing with these blockages to transcription, only some of which are understood. We examined the fate of RNA polymerase stalled by DNA-protein crosslinks (DPCs), as well as elongation inhibitors Streptolydigin (Stl) and Actinomycin D (ActD).

Our lab previously showed the importance of the tmRNA system for survival during DPC-formation, implying that transcription and translation are blocked by DPCs. Using 5-azacytidine-cytosine methyltransferase crosslinks as a model system for DPCs in *E. coli*, we tested knockout mutants of factors known to affect transcription using cell growth assays. Of these mutants, only *dksA* mutants were hypersensitive. However, western blots for tmRNA tagging showed that *dksA* mutants have increased rather than decreased tmRNA tagging, indicating that another unknown factor is responsible for enabling tmRNA activity. We also used the same cell growth assay to look for potential repair pathways for DPCs and found that *dnaK* knockouts were slightly resistant to DPCs while *dnaJ* knockous are sensitive. We propose a potential DnaK-independent role for DnaJ in DPC repair.

To isolate the effects of transcription elongation stalling, we treated cells with the elongation inhibitors Stl and ActD. Previous *in vivo* studies implied that Stl-inhibited polymerases are released from the DNA transcript via an unknown release factor. Using cell growth assays, Western blots for tmRNA tagging, and *in vitro* studies, we showed the transcription-coupled repair (TCR) factor Mfd is responsible for releasing both Stl- and ActD-stalled RNAP. We also treated *rpoB* mutants with ActD and found several ActD resistant mutants, implying alterations to RNAP are sufficient to eliminate ActD inhibition.

The tmRNA western blots also implied that Mfd has termination abilities in wildtype cells, leading us to perform RNAseq analysis on *mfd* knockout and overexpressing cells. We found that global transcription patterns are changed by altering Mfd levels, thus allowing us to propose a novel transcription regulatory role for Mfd.

Our studies show that polymerases stalled by DPCs and by elongation inhibitors are resolved by different mechanisms, emphasizing the importance of understanding the different pathways invovled in transcription elongation clearing. We also show that inhibitors such as ActD are effective against cells overexpressing the TCR pathway, which could have potential implications for the treatment of platinum-resistant tumors that have elevated levels of TCR.

Dedication

I would like to dedicate this dissertation to my Grandpa Krasich, who was my true role model in life.

Contents

List of Tables

List of Figures

List of Abbreviations

1. 5-aza-dC: 5-aza-deoxycytidine
2. ActD: Actinomycin D
3. AdoMet: S-adenosylmethionine
4. AML: Acute myelogenous leukemia
5. AzaC: 5-azacytidine
6. Bcm: Bicyclomycin
7. Cam: Chloramphenicol
8. ChIP: Chromatin Immunoprecipitation
9. CML: Chronic myelogenous leukemia
10. Cyt: Cytosine
11. DPC: DNA- protein crosslinks
12. FIC: Fractional inhibitory concentration
13. LB: Luria–Bertani broth
14. M.EcoRII: EcoRII cytosine methyltransferase
15. M.HhaI: HhaI cytosine methyltransferase
16. MDS: Myelodysplastic syndromes
17. MTase: DNA-cytosine methyltransferase
18. NER: Nucleotide Excision Repair
19. OD: Optical Density
20. Rif: Rifampin
21. RNAP: RNA polymerase
22. Stl: Streptolydigin
23. TCR: Transcription-coupled repair

Acknowledgements

I would like to thank my advisor, Ken Kreuzer, for his guidance and support throughout my graduate studies. His willingness to let me explore new areas of research has helped me develop my love for science.

I would also like to thank my committee members, Dr. Arno Greenleaf, Dr. Meta Kuehn, Dr. Chris Nicchitta, and Dr. Pei Zhou for their valuable advice and guidance. Their contributions to this project have helped assist me through several technical and scientific challenges.

Finally, I would like to thank my family for their constant love and support both throughout graduate school and my entire life. They have always been there to help me through every challenge that I've faced and helped me keep perspective on the important values of life.

1. Background

The general purpose of this dissertation is to examine the mechanisms by which cells address stalled transcription. Transcription is the process by which genetic information is translated from DNA into RNA. Due to its essential nature, transcription is a highly regulated and discontinuous process. RNA polymerase (RNAP), the protein responsible for transcription, frequently stalls at natural pause sites, either in a temporary manner or in a longer, permanently arrested backtracked state. It has been well documented that these natural pauses during elongation lead to careful regulation of RNAP by allowing time for *trans*-acting elongation or termination factors to act on the stalled polymerase (1-3). RNAP can also encounter roadblocks in the form of DNA damage or exogenous transcription inhibitors.

Stalled RNAP is problematic for the cell as not only does deleteriously-stalled RNAP reduce the production of RNA and protein, but it also inhibits other enzymes that translocate along DNA. Understanding how cells respond to blocked transcription allows for a deeper understanding about this vital process. Furthermore, it provides insight into the mechanism of the elongation inhibitors themselves. This dissertation specifically focuses on the model organism *Escherichia coli*, which is a useful tool for understanding basic procecces that are directly applicable to higher, more complex organisms.

1.1 Stages of transcription

The bulk of transcription is carried out by core RNAP, a roughly 400 kDa enzyme comprised of five subunits: $\sigma_2\beta\beta'\omega$. This core unit is highly conserved across all domains of life, emphasizing its essential nature and making the study of bacterial transcription relevant for the understanding of the transcription of higher organisms (4-7). RNAP is generally shaped like a "crab claw", with the β and β' subunits making the pincers of the claw. The catalytic active site is located at the base of the "hinge" of the claw. Other key structural components include the RNA exit channel, where the nascent RNA chain exits, and the secondary channel that leads to the catalytic site (4).

Core RNAP requires the binding of σ factor to form the holoenzyme and recognize sequence specific promoters. The selective binding of σ factors is a key regulatory step in transcription, as different σ factors respond to various environmental and stress stimuli and initiate the transcription of relevant genes. *E. coli* has seven σ factors, with σ^{70}, which is responsible for general housekeeping transcription, being the most abundant. σ's primary role is to orchestrate various steps of initiation, including recognition of promoter DNA and melting of the DNA at the transcription start site (8), though σ factors may also have an influence on transcription elongation (9-11).

Initiation begins with the holoenzyme binding to the -35 and -10 sequences of the promoter in an open conformation and melting the DNA, starting with the -11 position.

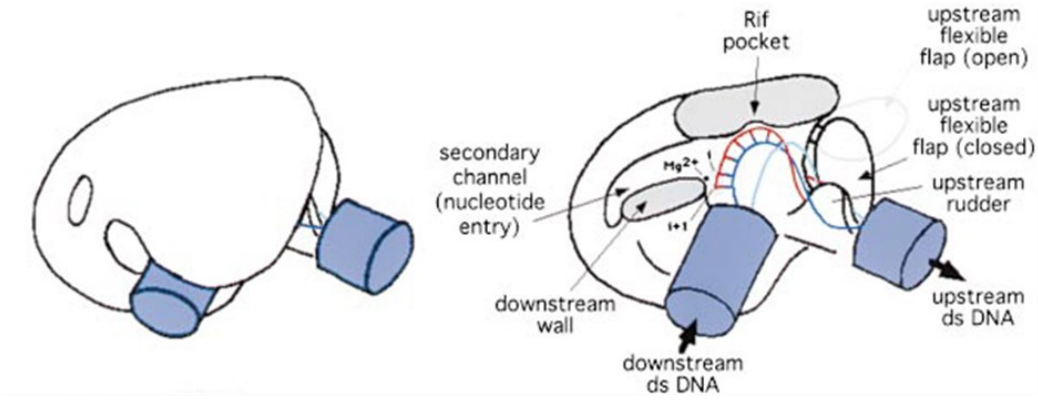


Figure 1: Structure of ternary RNAP complex

Schematic model of the structure of a ternary transcription complex. Double-stranded DNA is represented as blue cylinders. The DNA template strand is shown as a blue line; the nontemplate strand, a cyan line; the RNA transcript, a red line. Very little information is available to position the nontemplate DNA strand within the model; it is shown here for illustrative purposes only. (Left) View with intact RNAP molecule. (Bottom) Same view but with parts of the RNAP cut away (shown in gray) to reveal the inner workings of the complex, which are labeled (4).

RNAP then undergoes multiple rounds of abortive initiation, transcribing short transcripts that are rapidly released (12). The amount of promoter escape is dependent on promoter strength and regulation by initiation factors such as GreA and NusA (12-14) (for a list of all referenced genes and functions, see Appendix I). When the complex escapes past the first ~13-15 nt, allowing for a 7-9 bp RNA/DNA hybrid, RNAP undergoes a conformational shift into the closed complex to form the elongation complex (15). It is generally believed that σ is also released during the transition from initiation to elongation, although some researchers argue that σ is capable of traveling with or rebinding to the elongation complex (16-18).

Unlike initiation, elongation is a very stable and processive process. RNA polymerase can transcribe up to 80 nt/s, averaging 40 ntp/s *in vivo*, and can transcribe transcripts tens of kilobases long (19). The elongation complex (EC) is comprised of the DNA template, RNAP, the nascent RNA chain, and various elongation factors that travel with the EC (such as NusA) (20). RNAP forms three interactions with nucleic acids, leading to the stability of the EC: the HBS (RNA:DNA hybrid binding site), RBS ('tight' RNA-binding site), and UBS ('weak' upstream RNA-binding site) (21). During elongation, double-stranded DNA enters through the main channel until it is separated by fork loop 2 and the bridge helix, where non-template DNA exits past loop 2 and template DNA enters the catalytic active site. In the active site, template DNA and the

nascent RNA chain form an 8 ± 1 bp RNA-DNA hybrid, with the incoming NTPs entering the catalytic site though the secondary channel and the growing RNA chain exiting out the RNA exit channel. The template DNA then passes out of the RNAP and reanneals with the non-template strand ((15,22) and references therein).

To form a new transcript RNAP undergoes potentially thousands of rounds of the nucleotide addition cycle (NAC). During the NAC the RNA 3′ nt slides from the i + 1 site to the i site in the catalytic active site, opening up the NTP binding site for the incoming NTP. The incoming NTP first binds to the preinsertion site while RNAP is in an inactive "open" conformation, allowing for proper substrate recognition. RNAP then shifts into a "closed" conformation with the NTP transferring to the insertion site. Nucleotide addition is complete with the formation of the phosphodiester bond, and release of PPi (23-26).

Despite its processivity, transcription is highly discontinuous. Natural pause sites allow for essential regulation of transcription production and help to coordinate polymerase and ribosome activity in bacteria (27). Complexes can also be deleteriously stalled by a number of causes, such as DNA damage, exogenous RNAP inhibitors, and colliding replication forks (see below). Once stalled, elongation complexes can undergo a variety of fates including transcription restart, backtracking, or termination, depending on the cause of stalling.

In bacteria there are three stages of pausing: class I, class II, and ubiquitous pauses (28). Class 1 pauses are classified by the formation of RNA secondary structures, primarily hairpins. While the mechanism is still not fully understood, it is believed that the hairpin interacts with RNAP, stabilizing an active site rearrangement that temporarily halts the normal NAC (29). Class II hairpins are classified by RNAP backtracking. During backtracking, RNAP moves upstream along the DNA by various amounts, resulting in the removal of the 3' end of the nascent RNA from the catalytic center and extrusion of the tail through the secondary channel of RNAP. Under certain circumstances this pausing can lead to elongation arrest or recycling (30). Backtracked RNAP can also result in double-stranded breaks and genome instability if a replication fork collides (31). Ubiquitous pauses are the least understood form of pausing, with only a rough consensus sequence proposed and no other hallmarks of identification (32). Single molecule studies with ubiquitous pauses also showed they are not associated with backtracking (33).

Normal termination occurs in either a Rho-dependent or Rho-independent manner. Intrinsic termination occurs at sequence dependent sites, and account for termination at roughly 50% of annotated genes and 70% of noncoding RNA in *E. coli* (34). Intrinsic termination sites are comprised of a GC-rich palindromic element that forms a hairpin (comprised of a 8-9 bp stem and 4-8 nt loop) in free RNA, followed by an

6-8 bp oligo(T) sequence (T-stretch) (35). Elongating RNAP pauses at the unstable T-stretch, allowing time for the formation of the hairpin. Hairpin formation causes partial melting of the T-stretch, resulting in breaking the UBS and RBS contacts and weakening the HBS contacts. The transcription bubble is thus weakened and the DNA rewinds as RNAP is released (reviewed in (36)). In Rho-dependent termination, Rho recognizes and binds to C-rich *rut* (*R*ho *ut*ilization) sites on the nascent RNA chain. The RNA then translocates through Rho until there is little space between Rho and RNAP. This translocation requires there to be little to no impeding factors on the RNA, such as ribosomes or RNA secondary structures. Because the RNA translocates through Rho as RNAP is actively transcribing, RNAP is typically required to pause before Rho termination occurs, although backtracked and strong class I pauses are resistant to Rho-dependent termination. The exact mechanism of termination once Rho reaches RNAP is also still under debate, with current models including unraveling of the RNA:DNA hybrid due to tension caused by translocation and destabilizing conformational changes in RNAP due to allosteric binding of Rho (reviewed in (37)). Other factors are known to release RNAP when stalled at DNA damage or during replication-transcription collisions and are reviewed in section 1.4.

In addition to the inhibition of RNA production, failure of transcription to properly elongate or terminate can increase the number of collisions between the stalled

RNAP and other proteins translocating along the DNA. This is particularly problematic in the case of replication-transcription collisions, which potentially lead to replisome stalling and genome instability (31,38,39). Therefore, it is of importance to understand how the cell responds to and clears blockages in transcription. Our lab became interested in the fate of transcription complexes that were stalled in such a way that they could not be easily restarted, such as at a covalently bound protein or by a specific elongation inhibitor. This dissertation examines the fate of RNAP stalled by DNA-protein complexes and the elongation inhibitors Streptolydigin and Actinomycin D.

1.2 Physical block to RNAP: DNA-Protein Crosslinks

1.2.1. Causes and consequences of DNA-protein crosslinks

DNA-protein crosslinks (DPCs) occur when a protein becomes covalently bound to DNA. There are currently 4 classes of DPCs: the first three involve topoisomerases and DNA polymerases trapped at single-strand nicks and double-strand breaks, and the fourth involves proteins trapped on undisrupted double-stranded DNA (40). DPCs referred to in this document are Class IV DPCs. DPCs are caused by a number of agents, including ultraviolet light, radiation, metals and metalloids, various aldehydes, reactive oxygen species, and chemotherapeutic agents. Crosslinking mechanisms include direct covalent bond formation, chemical/drug linkers, or coordination with a

metal atom (41). Therefore, DPCs are chemically distinct and are influenced by a variety of factors. Due to the variable nature of DPCs, their consequences are difficult to interpret, but are known to include blockage of replication and transcription, formation of DNA double-stranded breaks, and cell death (42-44). In addition, DPCs have been implied to contribute to the cytotoxic, mutagenic, and carcinogenic effects of various chemical agents and chemotherapeutic drugs (44). Due to the wide range of types of DPCs and, therefore, the various causes and consequences, this background will be limited specifically to crosslinks induced by aza-C.

Aza-C is a cytidine analog with ribose sugar and a nitrogen atom at the carbon 5 (C5) position. In both mammalian and bacterial cells aza-C treatment leads to its incorporation into both RNA and DNA. A related compound, 5-aza-deoxycytidine (5-aza-dC), has a similar structure with a deoxyribose sugar, and thus incorporates primarily into DNA. aza-C and 5-aza-dC incorporates randomly into the genome, and incorporation into the inner C of CCWGG sequences leads to the covalent trapping of the EcoRII DNA-cytosine MTase (45) (see below for details and mechanism). This leads to a variety of effects including decreased MTase activity (44), replication fork blockage (42) , and transcription blockage (46). aza-C treatment also leads to inhibition of protein synthesis, although this is most likely due to its incorporation into RNA. Incorporation into tRNA inhibits tRNA MTases and thus tRNA methylation and processing, most

likely leading to the effect on protein synthesis (44) . Ultimately, aza-C treatment leads to decreased cell growth at low doses and cell death at high doses, although this effect is dependent on the presence of a cytosine MTase (47). To corroborate this, it has been shown that the level of cellular resistance to aza-C and 5-aza-dC is inversely related to active MTase levels (44) and the presence of a MTase overexpression plasmid dramatically increases cell sensitivity to aza-C (47) .

1.2.2. Mechanism of MTase binding to cytidine and 5-azacytidine

In normal DNA methylation, 5-methylcytosine is formed by the transfer of a methyl group from S-adenosylmethionine (AdoMet) to C5 of the target cytosine in a reaction catalyzed by cytosine-5-MTases. A highly conserved catalytic cysteine of MTase forms a covalent bond with C6 of the target cytosine, which causes a shift of electrons to C5 and subsequent attack on the donor methyl group from AdoMet. The remaining proton on C5 is then abstracted followed by β elimination, which reforms the double-bond between C5 and C6 and releases the enzyme (44).

In 1994 Klimasauskas *et al.* crystallized a ternary structure between a 13-mer oligonucleotide containing methylated 5-flourocytosine, the *HhaI* cytosine MTase (M.HhaI) and the reaction product S-adenosyl-L-homocysteine. They found that upon recognition of the target cytosine (Cyt) is flipped completely out of the DNA helix into the active site of M. HhaI. This "flipped-out" conformation aligns Cyt with the bound

AdoMet, and the conserved active-site loop in M.HhaI undergoes a large conformation shift towards the DNA binding cleft. The sulfhydryl group of the catalytic Cys-81, which performs the nucleophilic attack on Cyt, is brought to close proximity of Cyt as a result of the conformation shift (48). Further work by Klimasauskas *et al* showed that when M.HhaI binds to the recognition sequence it is in dynamic equilibrium with a variety of "open" and "closed" conformations. Subsequent binding of AdoMet then facilitates the stabilization of the "closed" conformation where Cyt is locked in the active site of M.HhaI. At this time the active cytosine performs its nucleophilic attack, causing a shift of electrons that results in C6 taking on sp^3 character. This facilitates attack of C5 on AdoMet, resulting in methyl transfer (49). This methyl transfer initiates the release of M.HhaI. First, the presence of the methyl group on C5 puts steric tension on Pro80 of M.HhaI, which is proposed to initiate a shift in equilibrium from the "closed" conformation to the "open" conformation of the enzyme. In addition, the presence of two water molecules located 4.4 Å and 4.2 Å from C5 provides general bases to facilitate the elimination of a proton from C5, resulting in the reformation of the double bond between C5 and C6 and the release of the enzyme (50).

As previously mentioned, aza-C has a nitrogen at the C6 position, which drastically alters the chemistry of the MTase reaction. First, unlike the reaction with normal Cyt, a stable covalent interaction between DNA-cytosine MTases can form in the

presence or absence of Ado-Met. This led Santi *et al.* to propose an AdoMet independent reaction mechanism: base recognition and subsequent MTase (M.HhaII was the MTase used in their studies) binding occurs via the formation of a covalent bond between the conserved catalytic cysteine residue of MTase and C5 of aza-C. This reaction then leads to the protonation of N6 in aza-C (45). This proposal would suggest that methylation of aza-C is not necessary for covalent bond formation; however, there is some disagreement over whether or not the methylation occurs. In support of Santi's proposal, Brank and Christman found that transfer of radiolabeled methyl groups from AdoMet to aza-C-containing oligonucleotides by M.HhaI was barely detectable (44). In contrast, Gabbara and Bhagwat were able to detect such transfer using the *EcoRII* MTase (M.EcoRII), leading to a revised proposal that in the presence of AdoMet N5 is methylated, but that an AdoMet-independent covalent bond could form in the absence of AdoMet (51). In either case, the presence of a nitrogen at the 5 position eliminates the possibility for the subsequent β-elimination step, blocking the formation of the 5,6 double-bond and preventing the release of the enzyme. In addition, Brank and Christman also tested the mobility of M.HhaI:aza-C-oligonucleotide complexes and found that in the presence or absence of AdoMet M.HhaI was always in the "closed" conformation, regardless of temperature (44). This again supports the original proposal by Santi *et al.* that aza-C inhibits DNA-cytosine MTases by covalently trapping the

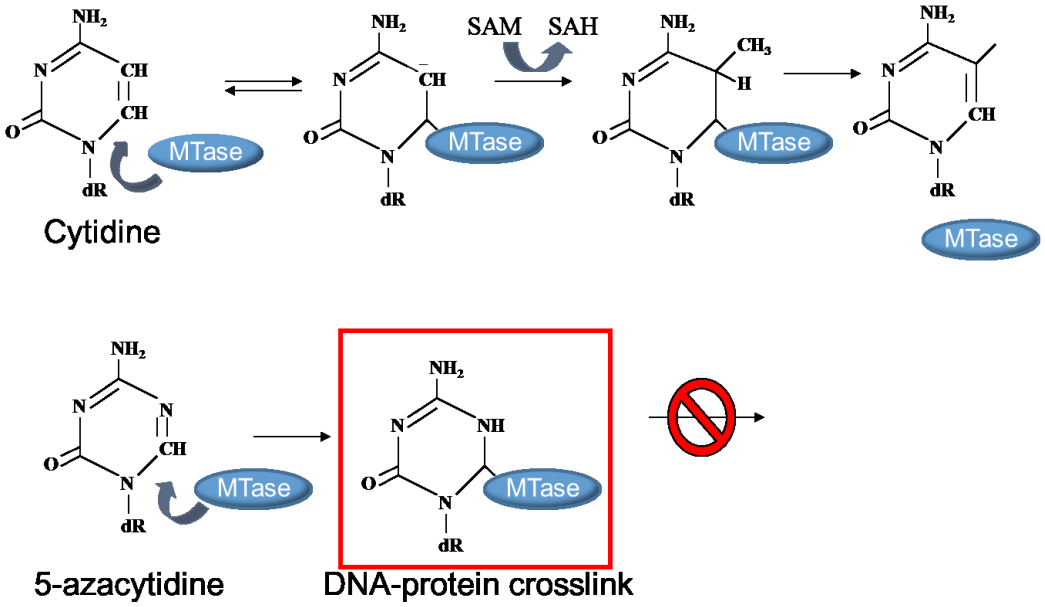


Figure 2: Methyl transfer reaction onto cytidine and 5-azacytidine
In normal DNA methylation, a highly conserved catalytic cysteine of MTase forms a covalent bond with C6 of the target cytosine, which causes a shift of electrons to C5 and subsequent attack on the donor methyl group from AdoMet. The remaining proton on C5 is then abstracted followed by β elimination, which reforms the double-bond between C5 and C6 and releases the enzyme (44). With aza-C, the nitrogen at the 5 position prevents the β elimination and the enzyme is covalently trapped.

enzyme at recognition sites. These complexes are extremely stable; they last up to 2-3 weeks when the complex is formed in the presence of AdoMet (44), and are stable to denaturation by Na-DodSO_4 (45), DNaseI digestion (52), 6 M urea (52), and precipitation by ethanol (52).

1.2.3. Significance of 5-azacytidine study

5-azacytidine-induced crosslinks are intriguing to study for multiple reasons. First, as previously mentioned, aza-C can serve as a model system for DPC study. Our lab uses the *EcoRII* MTase in our bacterial studies, which is a 60kDa protein with a recognition sequence of CCWGG. This becomes critical as different types of DPCs not only have different consequences, but the way the cell processes them varies as well. For example, Nakano and colleagues have shown that the size of the protein can affect the way that DPCs are processed: proteins 12-14 kDa and smaller can be processed by nucleotide excision repair (NER), while DPCs formed by larger proteins are potentially processed by homologous recombination (40); however, DPC processing is even less clear cut than this, as NER efficiency varies with the size of protein (53). Knowing the sequence that MTase targets also adds a technical advantage for directed studies.

In addition to its role in DPC study, aza-C has clinical implications as well. aza-C and the related compound 5-aza-deoxycytidine are FDA-approved hypomethylating agents for the treatment of myelodysplastic syndromes (MDS), in particular for

intermediate and high risk MDS (54). They have also been shown to be effective against acute and chronic myelogenous leukemia (AML and CML, respectively) (55). There are more than 10,000 cases of MDS diagnosed in the U.S. per year, with a 3-year survival rate of 35%; in addition, high-risk MDS has an average survival rate of 0.4-1.2 years with a high risk for development of AML. However, aza-C treatment has been shown to increase overall response rate (measured as the combination of partial and complete remission), overall survival, and prolong transformation to AML all at statistically significant levels. Prior to the use of hypomethylating agents cytotoxic therapy or supportive care were the primary means of treatment. Both of these cause potentially fatal secondary complications, whereas hypomethylating agents have a favorable side-effect profile (54).

Clearly aza-C and 5-aza-dC have vast potential for clinical significance. However, despite more than 40 years of research the mechanisms behind aza-C's cytotoxic and clinical activities are still not fully understood. Cytosine methylation affects mammalian cell expression (44) and it is suspected that reactivation of silenced tumor suppressor genes leads to the anticancer effect of hypomethylating agents (56); however, the aza-C-MTase adducts can also have a negative effect, such as proposed toxicity to bone marrow cells (56). The complexity of clinical consequences of aza-C treatment makes it imperative to understand the full consequences of aza-C

incorporation at a biochemical level so that the full potential of aza-C and related compounds as anticancer agents can be maximized.

1.3 Exogenous elongation inhibitors

1.3.1 Streptolydigin

Streptolydigin (Stl) is a transcription elongation inhibitor that inhibits RNA chain elongation (57-59). It is comprised of a streptolol moiety, a bulky tetramic acid group, and a sugar group (Figure 3). Unlike typical DNA binding elongation inhibitors (see below) Stl binds directly to RNAP. The Stl binding site in RNAP is ~20 Å from the active site and forms two main interactions: the streptolol moiety binds hydrophobically to the N-terminal portion of the bridge helix on the β′ subunit, and the tetramic acid group interacts with the central portion of the bridge helix and the trigger loop of the β subunit. These interactions stabilize the straight-bridge-helix RNAP-active-center conformational state (60,61). Crystal structures with RNAP and Stl and apo-holoenzyme show that the bridge helix is in a straight conformation when Stl binds, indicating that RNAP is most likely in the "open" conformation (61).

The mechanism of Stl inhibition of elongation is thought to be the stabilization of the bridge helix and the trigger loop. This stabilization prevents normal conformational

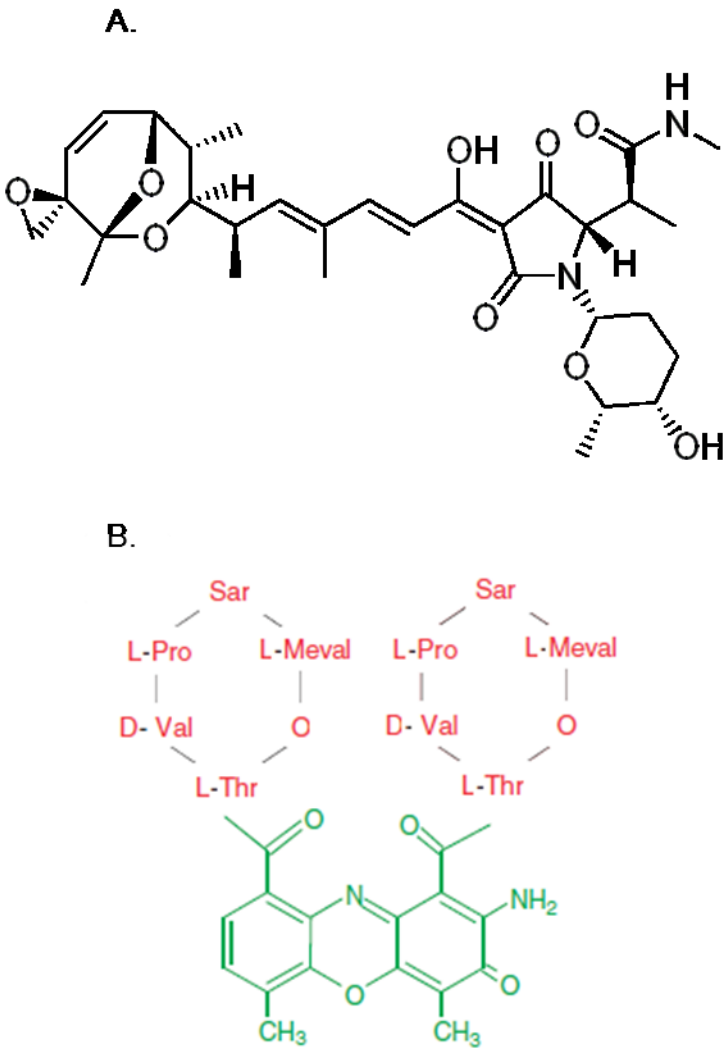


Figure 3: Structure of Stretpolydigin and Actinomycin D
Panel A is the chemical structure of Streptolydigin. Stl consists of a streptolol moiety, a bulky tetramic acid group, and a sugar group. Panel B is the structure for Actinomycin D. The planar phenoxazone ring system is shown in green and the pentapeptide side chains shown in red (62).

changes necessary for translocation of the growing nascent chain, effectively freezing RNAP in an inactive conformation (60,61). Streptolydigin does not appear to alter nucleotide binding to RNAP, phosphodiester bond formation, or DNA translocation (61). Elongation complexes stalled by Stl are very stable *in vitro;* complexes are still bound to DNA after 50 minutes of high Streptolydigin treatment, and removal of Stl results in resumption of transcription after a delay (59). However, in striking contrast, Stl-stalled elongation complexes are irreversibly destabilized *in vivo*, presumably due to being released by some unidentified release factor (58).

1.3.2. Actinomycin D

Actinomycin D (ActD) is a small molecule comprised of a planar tricyclic phenoxazone ring and two cyclic pentapeptide side chains with antibiotic and anticancer capabilities (62,63). It is primarily known as a DNA intercalator, with the ability to intercalate into double stranded DNA, bind to single-stranded DNA, and hemi-intercalate into single stranded DNA. Crystal structures have shown that phenoxazone ring system intercalates preferentially between dinucleotide sequences of dGpC. The pentapeptide chains lie in the narrow groove of the structure and form hydrogen bonds with guanine residues on the chain opposite of ActD intercalation (64). A more recent force spectroscopy study showed the ActD can intercalate into stable B-

DNA with very slow on and off-rates. Destabilization (in the form of DNA stretching) increasing the binding stability 20-fold , the on rate of intercalation by 100-fold, and the off rate of release 5 fold (62).

Despite years of research, the exact mechanism by which ActD carries out its antibiotic and anticancer capabilities is still under debate. Early studies found that ActD preferentially bound to distorted double helical DNA, particularly β-DNA found in transcription bubbles, and proposed that binding of ActD to the leading edge of the transcription bubble effectively "pins" RNAP in place and inhibits transcription elongation (64). The authors of the force spectroscopy study supported this model and proposed that the increased transcriptional activity of tumors is what makes ActD so effective. Other models propose that ActD binds to B-DNA, stabilizing double-stranded DNA and blocking replication forks (65). To the best of our knowledge, direct interaction of ActD with RNA or DNA polymerases has not been addressed.

1.4 Fate of paused or stalled RNAP

A common way the cell recycles stalled elongation complexes is through the activity of TRCF (transcription repair coupling factor), also known as Mfd (mutation frequency decline). Mfd is a member of the helicase superfamily 2, and has ATPase, translocase, and dsDNA binding activity but not strand-separating helicase activity (66).

Mfd couples transcription termination and DNA repair by recycling stalled polymerase and recruiting nucleotide excision repair factors to the site of the DNA lesion. Mechanistically, Mfd is recruited to a stalled elongation complex through its RNAP binding domain D4, binding directly upstream of the transcription bubble (30,67). Subsequent DNA binding leads to a conformational shift, causing movement of the helical hairpin into the dsDNA binding cleft. This hairpin serves as a ratchet, pushing RNAP forward and DNA backwards. Depending on sequence and RNAP conformation, Mfd then either enhances forward translocation or removes the stalled RNAP from the DNA. For instance, if RNAP is backtracked, MFD uses its ATP dependent translocase activity to push forward backtracked RNAP. This concomitantly rewinds the transcription bubble from upstream and unwinds the RNA/DNA hybrid (67,68). When there is a physical obstruction in the way, such as a DNA-bound protein or DNA lesion, this ratcheting motion is proposed to dislodge RNAP from DNA (66,69). If the block was induced by DNA damage, once RNAP has been cleared from the DNA UvrA is recruited to the site via Mfd's UvrA binding site in the N-terminal UvrB homology domains (66,67) to begin the process of nucleotide excision repair.

In addition to Mfd, bacterial cells have other pathways to resolve replication/transcription complexes stalled at protein roadblocks or other blocking lesions. The DinG, UvrD, and Rep helicases have been implicated in preventing or

mitigating the damage from collisions between the replication machinery and bound proteins (such as RNA polymerase), and at least Rep and UvrD have protein removal activity *in vitro* (70-73). Transcription terminator Rho has been shown to prevent double stranded DNA breaks, presumably by removing RNAP ahead of the replisome and preventing damaging collisions (74). It also has been shown to remove polymerases stalled at a tightly bound protein *in vitro* (75). Another transcription factor, HepA, has been shown to activate transcription by recycling RNAP, and potentially plays a role during DNA damage (76,77).

In some cases the transcription blocked by a lesion may be able to bypass the lesion altogether, through the assistance of transcription factors or trailing RNA polymerases in the case of transcription. GreA and GreB are elongation factors that travel with the transcription complex, and have been shown to induce cleavage of the 3' proximal dinucleotide from the nascent RNA by RNAP, allowing for restart of transcription at the new 3' end (78,79). GreA and GreB also stimulate activation of backtracked elongation complexes (80). DksA, along with ppGpp, has numerous effects on elongation complexes and has been shown to prevent replication/transcription collisions (73,81). Trailing RNA polymerases have also been shown to help push stalled elongation complexes past roadblocks (82).

1.5 Major questions to be addressed

The purpose of my dissertation research is to understand the molecular response to transcription inhibition with the downstream application of enhancing cancer and antibiotic therapies. Due to the variety of different fates and cellular consequences that stalled elongation complexes might have, this dissertation focuses specifically on the case where transcription cannot be easily restarted with the aid of elongation factors. To model this system, we look at three different inhibitors: DNA-protein crosslinks, Streptolydigin, and Actinomycin D.

As mentioned above, it has been shown that a specific form of DNA damage known as DNA-protein crosslinks (DPCs) lead to blockage of elongating transcription complexes *in vitro,* but to date the fate of these transcription complexes *in vivo* remains unknown. In the first part of this work, we utilize crosslinks formed between aza-C and cytosine MTase as our DPC model system. The main questions we ask are: 1) What happens to transcription blocked at a DPC and 2) How do DPCs get repaired (as this might influence the fate of already-stalled transcription complexes)? Our lab previously conducted a transposon mutagenesis screen looking for Aza-C hypersensitive mutants and identified 24 mutations in 16 genes (83). One gene was *ssrA*, which encodes the tmRNA gene product in *E. coli*. tmRNA is responsible for recycling stalled ribosomes and tagging the resulting truncated peptides for degradation. Since transcription and

translation are coupled in bacteria, the involvement of the tmRNA system implies that transcription is blocked as well. tmRNA binds to the A-site of the ribosome, and a nascent RNA chain that is still attached to a translating ribosome would occlude the A-site. Therefore a third question we address is how tmRNA accesses the stalled ribosome.

The transcription elongation inhibitor Stl binds directly to RNAP and traps it in an inactive conformation that cannot be restarted by simply being "pushed forward". Previous work showed that Stl-inhibited polymerases are released from the DNA transcript *in vivo*, and it is likely that transcription complexes blocked at DPCs are also released (58). ActD is a eukaryotic transcription elongation inhibitor that is known to intercalate into DNA, but the specific mechanism behind RNAP inhibition in bacteria is still unknown. Therefore the main questions we address in this section are: 1) What is the release factor responsible for clearing Stl-stalled RNAP and 2) does the same release factor(s) work for ActD treatment? Because so little is known about the exact mechanism of ActD inhibition of transcription, we will also be looking for RNAP mutants that are resistant to ActD treatment.

2. Functions that protect *Escherichia coli* from DNA-protein crosslinks

Portions of this chapter have been published in Molecular Microbiology (84). The transposon mutagenesis screen mentioned in this section was conducted by Sunny Wu and Kenny Kuo.

2.1 Introduction

In this Chapter, we aim to understand the effects of DNA-protein crosslinks (DPC) on transcription elongation. DPCs are a form of DNA damage that occur when proteins become aberrantly covalently attached to DNA. They can be caused by a number of factors such as formaldehyde, UV and IR treatment, metals and metalloids, reactive oxygen species, and chemotherapeutic agents. The variety of causes as well as the types and locations of proteins that form DPCs makes them very difficult to study, and as a result DPCs are the least understood form of DNA damage. To combat the challenges of DPC study, our lab uses 5-azacytidine (aza-C), a cytidine analog that covalently traps cytosine methlytransferases at specific recognition sites, as a model system.

It has been shown that DPCs have a number of cellular consequences including blocking replication, transcription (*in vitro)*, and recombination (42-44). We have also previously reported that the tmRNA system is required for cell survival during DPC formation, implying translation is alo blocked (84). We report here that the tmRNA

system is required to recycle the stalled ribosomes. The first step in tmRNA-mediated ribosome recycling is tmRNA binding to the empty A-site of a ribosome, which leads to the question of how tmRNA is able to access the stalled ribosomes if they are still on the nascent transcript. We were unable to identify any RNase mutants that were hypersensitive to aza-C, implying that A-site cleavage is not responsible for tmRNA tagging. A main focus of this chapter is thus to identify potential RNAP releasing proteins. Since the repair of DPCs might be coupled to the clearing of the blocked transcription complexes, we also look for potential repair pathways that lead to the clearing of DPCs.

2.2 Results

Requirement of the tmRNA system during DPC formation implies transcription blockage *in vivo*

Our lab previously conducted a transposon mutagenesis screen for aza-C hypersensitive mutants in an attempt to identify proteins involved in the repair and tolerance of DPCs. This screen resulted in the identification of a total of 24 insertion mutations, localized to 16 different genes, qualified as aza-C hypersensitive (83). This screen both confirmed genes already identified as aza-C hypersensitive and isolated new ones. One of the unique genes isolated was the *ssrA* gene, which encodes the tmRNA gene in *E. coli* (84). Because transcription and translation happen concurrently in bacteria, the requirement of the tmRNA system – which is known to recycle stalled

ribosomes – for cell survival implies both translation and transcription are blocked during DPC formation. Knockouts of *smpB*, an essential cofactor to tmRNA, were also shown to be hypersensitive, further confirming involvement of the tmRNA system (84).

In addition to recycling stalled ribosomes, tmRNA also adds a degradation tag to the truncated peptides during trans-translation and prevents accumulation of unproductive products (85). To separate the different functions of tmRNA and distinguish which is responsible for cell survival during DPC formation, we used the mutant *ssrA-H*$_6$ gene, which encodes a tmRNA with a hexahistidine stretch of codons in place of the coding sequence for a functional proteolysis signal (86). This mutant tmRNA is functional for dissociating blocked ribosome complexes but causes a gross deficiency in downstream proteolysis. We introduced a p15A-derived plasmid with the *ssrA-H*$_6$ mutant, wild-type *ssrA* or empty vector into *ssrA* knockout cells. Due to plasmid compatibility issues, we used an M.EcoRII expression plasmid, pR234, in which the methyltransferase expression is under control of a P_{tac} promoter (87). Expression from this plasmid is leaky in the absence of IPTG, and so we also generated a set of strains in which the M.EcoRII coding sequence had been deleted from the expression plasmid to use as negative controls. The *ssrA* deletion cells expressing M.EcoRII but no plasmid-borne tmRNA showed strong sensitivity to aza-C, while the cells with the

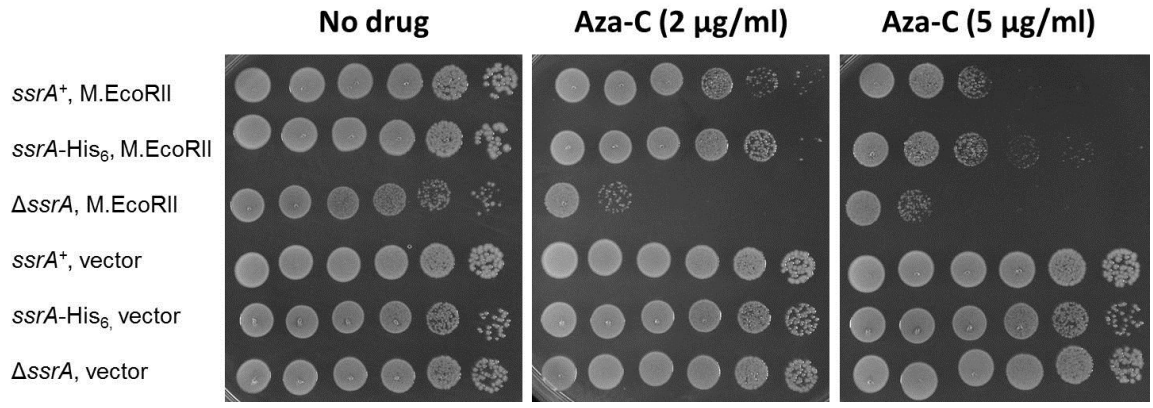


Figure 4: tmRNA is required to recycle ribosomes.
Aza-C hypersensitivity is relieved by the *ssrA-H6* mutant. Overnight cultures of strain HK22 *ssrA::kan* derivatives carrying two plasmids were diluted to 4 x 10^8 cells/ml. The identities of the plasmids are indicated on the left: ssrA$^+$ (pKW11); ssrA-H6 (pKW24); *ssrA* (pKW1); M.EcoRII (pR234); vector (pRK1). Ten-fold serial dilutions were generated across a microtitre plate and 5 ml of each dilution was spotted onto LB plates with no aza-C (left panels), aza-C at 2 mg/mL (middle panel), or aza-C at 5 mg/ml (right panel). Plates were photographed after overnight incubation at 37°C.

mutant *ssrA-H*$_6$ gene were resistant (Fig. 4). This implies that the ribosome recycling ability of tmRNA, not the degradation of the stalled peptides, is required for survival.

RNase activity is not responsible for creating tmRNA substrates

In the well-studied tmRNA pathway, the first step of ribosome release involves tmRNA binding to the empty A-site of the ribosome (85). For this to occur, the ribosome normally translates to the end of a truncated mRNA transcript (lacking a stop codon), or the transcript is cleaved to generate an empty A-site (88,89). In the case of DPCs and blocked transcription/translation complexes, how does tmRNA access the stalled ribosomes?

One model is that a ribonuclease cleaves the nascent transcript near the A-site of a stalled translation complex, generating a truncated transcript and a premature 3' mRNA end for any upstream ribosomes on the same transcript. If A-site cleavage is critical for tmRNA function at DPC-blocked transcription/translation complexes, we would expect to observe aza-C hypersensitivity in cells that are deficient in RNase II, which is implicated in A-site cleavage (see (90) and Discussion). For this and all subsequent drug sensitivity profiles, sensitivity levels were conducted by measuring growth curves in the presence of varying concentrations of aza-C in 96-well plates using a Biotek ELx808™ Absorbance Microplate Reader (which allows incubation at 37°C and agitation for aeration). Sensitivity levels were compared by processing the growth curves in a manner that corrects for differences in growth rates. For each mutant and the wild type, we first determined the growth rate of the drug-free culture during its exponential phase (OD_{560} values from 0.01 to 0.1; in every case, these data points

matched a simple exponential curve with R^2 values of greater than 0.99). We next determined the first time point at which the growth rate had fallen to half that value (or less), and plotted the accumulated cell density (as a fraction of the drug-free value) at that time point for the various drug concentrations. Using this test, we found no different in sensitivity levels between wildtype and *rnaseII* mutants (Figure 5). These results argue that A-site cleavage, at least by the most common pathway, is not necessary for tmRNA activity in this situation. We also tested mutants lacking other exonucleases with known roles in tRNA processing (RNase D, PH, and LS). However, knockout cells for each of these were no more sensitive to aza-C treatment than wild-type cells (data not shown).

Candidate approach: Functions potentially involved in clearing DPC-stalled transcription/translation complexes

An alternative model to A-site cleavage for tmRNA activity is that RNAP is released before tmRNA activity can occur. RNAP release would lead to the release of the nascent transcript, allowing translation to continue to the end of the truncated transcript and resulting in an empty A-site. Therefore we attempted to identify the release factor responsible for creating substrates for the tmRNA system. A well-studied mechanism for releasing RNA polymerase blocked by DNA damage (e.g. UV-induced

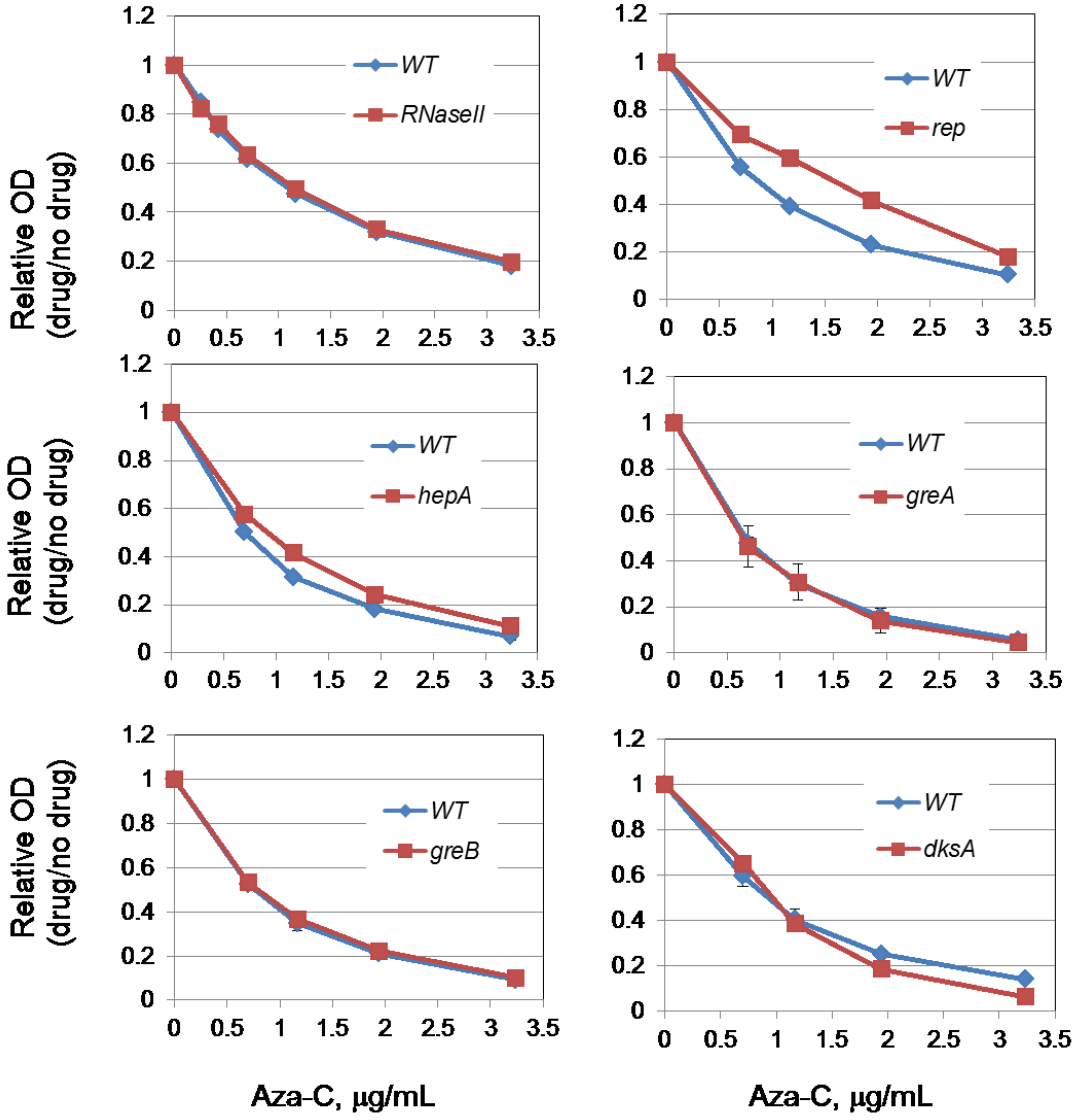


Figure 5: Aza-C hypersensitivity profiles of potential transcription altering knockout strains.

HK22 (WT) or mutant derivatives (containing pBAD-MEcoRII) were tested for aza-C sensitivity using a microtiter plate sensitivity assay. A standardized point in the growth curve was first determined for each cell line, namely the time at which the growth rate of the no-drug culture dropped to 50% of the earlier exponential rate. The OD_{630} value at this time was then taken for each drug concentration and divided by the no-drug control to account for differences in growth rate. Three independent cultures were measured for each strain, with the exception of the *rep* mutant, where 6 independent cultures were grown and measured due to the poor growth and slight variability of this strain. The error bars indicate the standard deviation for these cultures.

pyrimidine dimers) involves the Mfd helicase, which also couples transcriptional blockage to excision repair (66,91). Mfd therefore is a likely candidate for releasing DPC-stalled transcription complexes. Strikingly, a previous member of our lab found the *mfd* single mutant was no more sensitive than the wild-type, strongly arguing against a role of transcription-coupled repair during DPC formation.

Termination factor Rho is involved in terminating transcription at many natural terminators, and also induces termination when translation is blocked due to an upstream nonsense mutation (transcriptional polarity) (3). Furthermore, Rho protein induces termination and release of RNA polymerase blocked *in vitro* by a tightly bound protein (75). These results suggest Rho is a reasonable candidate for the alternative polymerase release function. In order to test the possible involvement of Rho, we used the Rho-specific inhibitor bicyclomycin (92). If Rho is involved in releasing RNA polymerase blocked by aza-C-induced DPCs we predicted that bicyclomycin would be synergistic with aza-C for growth inhibition. We used cells that express M.EcoRII from the pBAD-MEcoRII plasmid, and prepared microtitre plates with a double drug (aza-C and bicyclomycin) serial dilution in checkerboard fashion, thereby testing numerous combinations of drug concentrations for growth inhibition.

To assess whether or not the drugs act synergistically, we processed the data in two different manners. First, we looked for synergy using a graphical representation in which the two drug concentrations constitute the *x* and *y* axes, and the amounts

required to inhibit growth to 95%, 75% or 50% (isoboles) are plotted in different colors. Synergistic drug interactions are revealed by a concave shape to the isobolic curve, antagonistic interactions by a convex shape, and lack of drug interaction by a relatively straight line (93). The data from four separate experiments are all plotted in Figure 6, with the theoretical lines for no drug interaction shown as dashed lines (connecting the two experimentally determined values for each drug alone). The multiple data points fall quite near the theoretical line for no drug interaction, with no indication of a synergistic (concave line) relationship and a slight but unconvincing hint of an antagonistic (convex line) relationship.

The second method commonly used to assess synergy, supported by the American Society for Microbiology, is the so-called fractional inhibitory concentration (FIC) index ((94); also see Instructions to Authors for the ASM Journal Antimicrobial Agents and Chemotherapy). For each level of drug A, the FIC is calculated for the first concentration of drug B that gave the indicated level of inhibition, with FIC = [(MIC of drug A tested in combination) / (MIC of drug A tested alone)] + [(MIC of drug *B*-tested in combination) / (MIC of drug *B*-tested alone)]. The American Society for Microbiology recommends that synergy is supported by FIC index values less than 0.5, while antagonism is supported by FIC index values greater than 4 (an FIC of 0.5 could reflect a situation where 1/4 the concentration of each drug is required in combination to give the

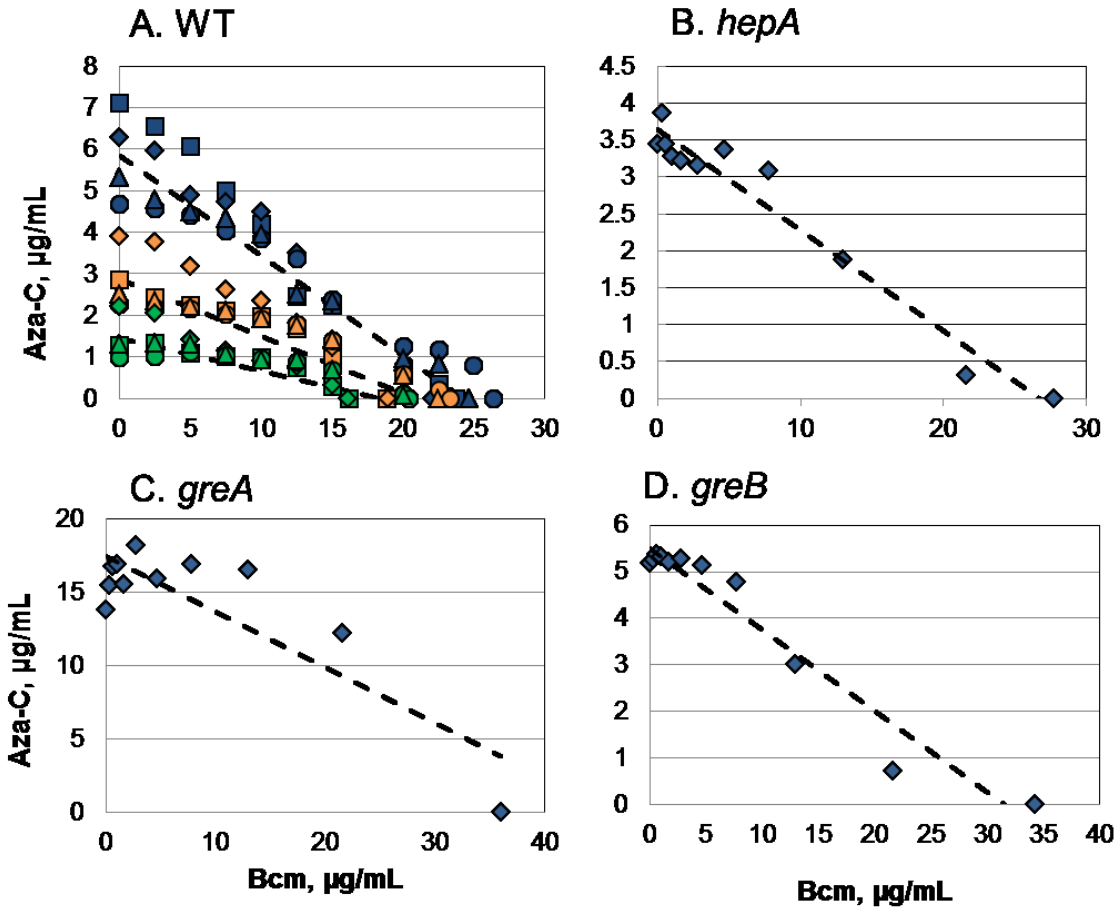


Figure 6: Isobolic tests for synergy with Rho inhibitor bicyclomycin
Growth curves were measured in each well of a 96-well microtitre plate, with varying concentrations of bicyclomycin (right to left) and aza-C. The strain for the bicyclomycin/aza-C experiment were (A) HK22 pBAD-MEcoRII (B) *hepA*, (C) *greA* and (D) *greB*. A detailed description of the data analysis and processing are presented in Methods. Briefly, at each concentration of bicyclomycin (Bcm), the concentration of aza-C necessary to inhibit growth by 95% (blue), 75% (gold) or 50% (green) was estimated. In addition, the concentration of bicyclomycin necessary for those levels of growth inhibition (in the absence of the second drug) was estimated from the bicyclomycin inhibition curve. The data from each of four repetitions (on different days) were plotted with different symbols (squares, diamonds, circles and triangles). The dashed lines connect the average determined MIC value of each drug alone.

Table 1: FIC values from the aza-C / bicyclomycin titrations.

95% Growth Inhibition

Experiment	Aza-C	Bcm	FIC
1	7.5	0	
1	0	25	
1	1	20	**0.93**
1	2.5	12.5	**0.83**
1	7.5	2.5	**1.10**
1	5	7.5	**0.97**
2	7.5	0	
2	0	22.5	
2	7.5	2.5	**1.11**
2	5	5	**0.89**
2	2.5	15	**1**
2	1	20	**1.02**
3	5	0	
3	0	27.5	
3	2.5	15	**1.05**
3	1	20	**0.93**
4	7.5	0	
4	0	25	
4	5	2.5	**0.77**
4	2.5	12.5	**0.83**
4	1	20	**0.93**
		Ave:	**0.95**
		St. Dev:	0.11

75% Growth Inhibition

Experiment	Aza-C	Bcm	FIC
1	5	0	
1	0	20	
1	2.5	7.5	**0.88**
1	1	15	**0.95**
2	5	0	
2	0	20	
2	2.5	10	**1.00**
2	1	20	**1.20**
3	2.5	0	
3	0	25	
3	1	20	**1.20**
4	2.5	0	
4	0	22.5	
4	1	10	**0.84**
		Ave:	**1.01**
		St. Dev:	0.16

50% Growth Inhibition

Experiment	Aza-C	Bcm	FIC
1	2.5	0	
1	0	20	
1	1	7.5	**0.78**
2	2.5	0	
2	0	20	
2	1	10	**0.90**
3	1	0	
3	0	22.5	
4	2.5	0	
4	0	22.5	
4	1	10	**0.84**
		Ave:	**0.84**
		St. Dev:	0.06

From the OD630 measurements at the set end point the fractional inhibitory concentrations (FIC) were calculated. For each level of aza-C, the FIC was calculated for the first concentration of bicyclomycin (Bcm) that gave the indicated level of inhibition (95%, 75% or 50%), and for each concentration of Bcm, the FIC was calculated for the first concentration of aza-C that gave 95%, 75% or 50% inhibition. This Table summarizes these FIC values from four different experiments for the three different levels of growth inhibition. The MIC values for each drug alone in each experiment are also shown (with no FIC value). The average and standard deviations from all FIC values are shown at the bottom.

same growth inhibition as each drug alone at 1x concentration; the theoretical value for no interaction whatsoever is 1.0). In the same tests shown graphically in Fig. 6, we found the following FIC index values: 0.95 (±0.11) for 95% inhibition; 1.01 (± 0.16) for 75% inhibition; and 0.84 (±0.06) for 50% inhibition (see Table 1). These FIC values clearly do not support a synergistic (or antagonistic) relationship between aza-C and bicyclomycin. With the caveat that multiple drug experiments need to be interpreted with caution, these results argue against an involvement of termination factor Rho in releasing RNA polymerase blocked by DPCs.

As mentioned previously, other factors are known to release RNAP in different stressful situations. Two candidates that arose in the transposon screen mentioned above, DinG and UvrD, along with the helicase Rep, were implicated as having overlapping roles to avoid transcription-replication conflicts, particularly when the collisions were from opposite directions (70). However, if either helicase was releasing the nascent transcript to allow access by the tmRNA system, we would expect decreased SsrA tagging in its absence. We found no such decrease in spite of the clear hypersensitivity of each mutant (data not shown).

The third helicase implicated in transcription-replication conflicts by Boubakri *et al.* (70), Rep, was not identified in the transposon screen above, but could have been missed since the screen was not saturating. Rep is of particular interest because it was recently shown to play a major role in minimizing replication pausing at sites of bound

protein, particularly transcription complexes, presumably by displacing the (non-covalently) bound protein (95). In the microtiter plate liquid growth assay, *rep* cells showed slight resistance to aza-C treatment (Figure 5 and Figure 7), however spot tests of *rep* cells on solid media showed no change in aza-C sensitivity compared to the wild type (Figure 8). The slight resistance in the microtiter plate assay could be an indirect effect of the relatively poor growth of the *rep* cells (see no-drug curves in Figure 7). Overall, these experiments provided no support for the model that accessory helicases release the transcript from DPC-stalled transcription complexes to allow access by the tmRNA system.

HepA is another factor known to be involved in RNAP recycling and we therefore tested *hepA* cells for aza-C sensitivity. Again we found that *hepA* cells were no more sensitive to aza-C than wild-type (Figure 5; perhaps even slightly resistant). Furthermore, treating *hepA* with bicyclomycin (Rho inhibitor) did not lead to sensitivity, arguing against redundant roles of Rho with HepA (Figure 6).

Since no potential release factor emerged from these tests, we turned to factors known to stimulate elongation past blocking lesions, namely GreA, GreB, and DksA. Strains carrying *greA* or *greB* mutations showed no change in sensitivity to aza-C, implying that these factors play no role in relieving DPC-induced elongation blockage (Figure 5). In addition, treating *greA* or *greB* cells with bicyclomycin also did not lead to aza-C hypersensitivity (Figure 6).

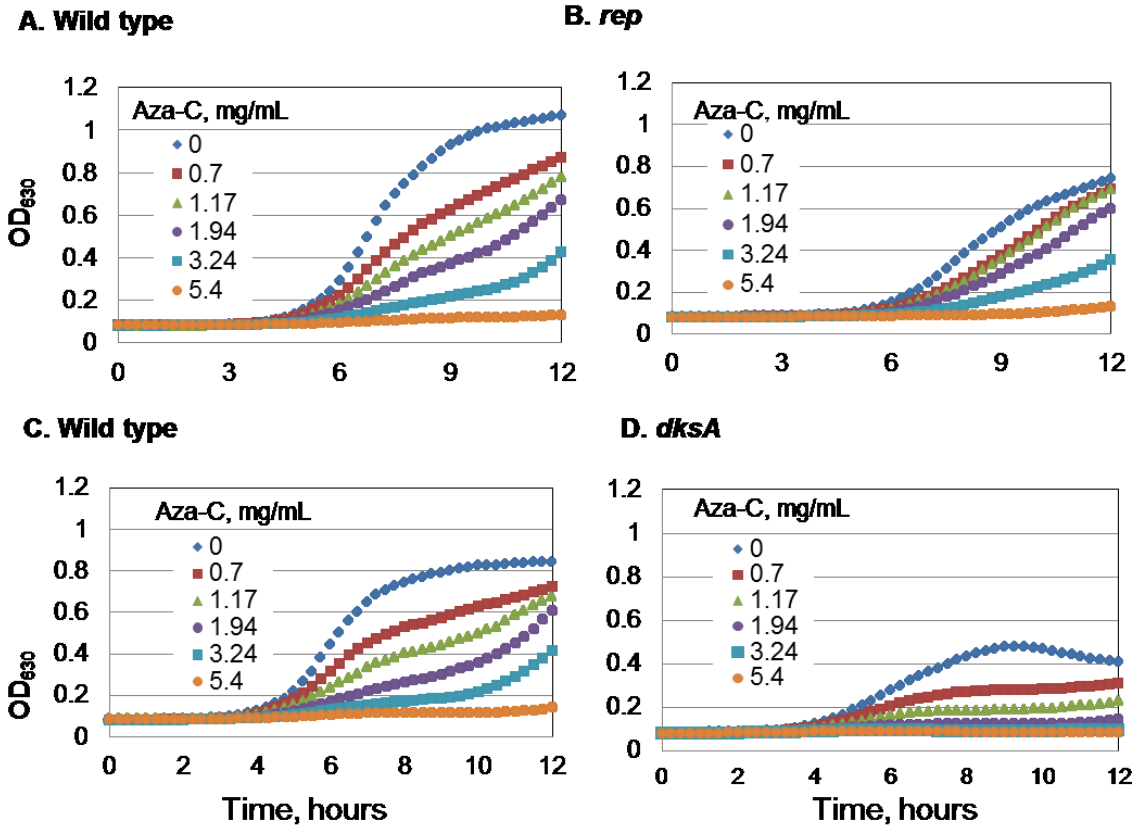


Figure 7: Aza-C hypersensitivity profiles for slow growing *rep* and *dksA* strains. Panels A and B are representative growth curves for HK22 (WT) and *rep* cells containing pBAD-MEcoRII, and panels C and D are representative growth curves for HK22 (WT) and *dksA* cells containing pBAD-MEcoRII. Overnight cultures were diluted to OD_{560} = 0.5, then diluted 1:2000 in LB with tetracycline and mixed with an equal volume (75 ◎l each) of LB containing the indicated amounts of aza-C in 96-well plates. Cells were grown for 12 hours at 37°C and cell turbidity was measured every 15 minutes. The data in Figures 5 and 9 for the *rep* and *dksA* cells were calculated from these growth curves.

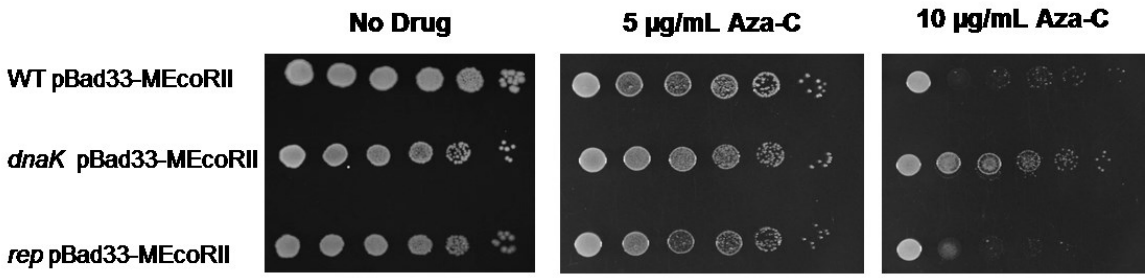


Figure 8: Spot tests for aza-C resistance of *rep* and *dnaK* strains.
HK22 *dnaK* and HK22 *rep* cells containing pBAD-MEcoRII were diluted from overnight culture to approximately 4 x 10^8 cells/ml. Ten-fold serial dilutions were generated across a 96-well plate and 5 µl of each dilution was spotted onto LB plates containing 5% arabinose and aza-C at 0, 5 or 10 µg/mL. Plates were photographed after overnight incubation at 37°C. Spot tests were conducted to either confirm resistance or to rule out resistance in the microtiter plates due to the slow growing nature of the mutants.

Under our standard sensitivity test conditions, the *dksA* knockout strain showed similar aza-C sensitivity as the wildtype (Figure 5). However, when the effect of aza-C was measured later in growth (when the growth rate of the no-drug control culture dropped to 10% of its maximal value), the *dksA* cells appeared modestly hypersensitive (Figure 9A). One caveat is that the very poor growth of the *dksA* mutant might somehow skew the sensitivity results (see the no drug curves in Figure 7). In addition, overexpressing DksA from the pCA24N plasmid lead to resistance of aza-C, even when measured at the 50% maximum growth rate time point (Figure 9B).

The simplest explanation for these growth results is that, under these conditions, DksA releases RNAP and the transcript from DPC-stalled complexes. If so, inactivation of DksA should decrease or eliminate SsrA tagging. We conducted Western blots using a primary antibody against the degradation tag of tmRNA. Truncated peptides that contain the degradation tag from tmRNA are typically degraded by proteases, including the ClpXP and Lon and proteases (96,97). HK22 cells are protease proficient, so any tagging that is detected is the result of the proteolysis system being overwhelmed.

However, Western blots for SsrA tagging levels (in late-growth cells) revealed slightly elevated levels of tagging induced by aza-C in *dksA* cells relative to wild-type cells (Figure 9C). This result argues against a role of DksA in releasing RNA polymerase and the transcript. The slight increase in tagging is consistent with a speculative alternative model, in which DksA helps RNAP bypass the DPC lesion without removing

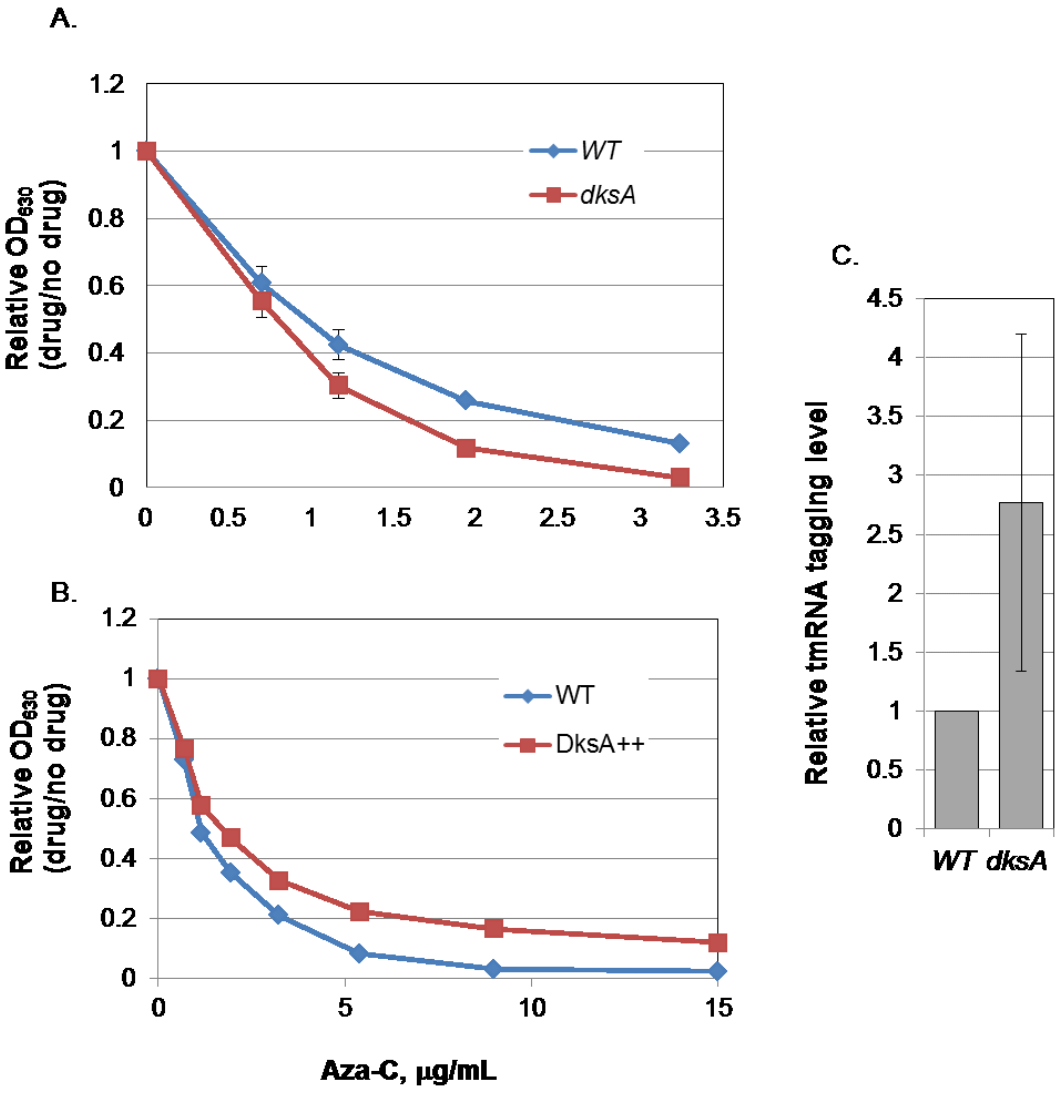


Figure 9: *dksA* knockouts are slightly hypersensitive to aza-C but do not have decreased tmRNA tagging.

Panel A is the comparative titration curves for the HK22 *dksA* strain shown in Figures 5 and 7, with the exception that the OD_{560} values plotted were from the time at which the growth rate of the no-drug culture dropped to 90% of the maximum exponential rate. Panel B is a comparative titration curve for cells overexpressing DksA activity. Panel C is the quantitation of western blots from extracts from HK22 (WT) or *dksA* cells with pBAD-MEcoRII using polyclonal antibodies to the degradation tag of tmRNA. HK22 cells are proficient in the protease systems, and accordingly peptide accumulation is due to oversaturation of those systems. Error bars for all panels are the standard deviations for three independent cultures.

it from the DNA, while some competing process releases the polymerase and transcript (see Discussion). *E. coli* RNAP can transcribe through reversibly bound proteins *in vitro* (82,98,99) and T7 RNAP can transcribe through a 44 kDa DPC *in vitro*, albeit with low frequency (100). While these results are consistent with a novel and interesting role for DksA in DPC biology, we were unable to identify any function required for tmRNA access after transcription/translation blockage by DPC (see Discussion).

Clarifications of genes identified in the aza-C hypersensitivity screen

As mentioned above, little is known about the exact repair mechanisms for DPCs, especially of bulky proteins like the EcoRII methyltransferase. Therefore another goal of our lab is to identify the protein(s) responsible for DPC repair. For the purposes of this dissertation we are also interested in the repair mechanism as it might shed some light on the fate of stalled transcription. If a transcription complex is stalled at a DPC it would be logical that the RNAP must be cleared first before repair can occur simply due to steric clashes; however, our inability to identify a RNAP clearing protein suggests that other mechanisms are possible.

The hypersensitivity screen identified 24 mutations in 16 genes which included genes involved in recombination, cell division, translation, tRNA modification and chaperone pathways (83). Exact location for each mutant was determined and is summarized in Table 2 (while previous lab members isolated the listed mutants, the

Table 2: Summary of transposon-insertion mutants hypersensitive to DPC inducer aza-C

Gene	9 bp repeat[a]	Last Nucleotide[b]	Function
dinG	CATTATTGT	834439	DNA helicase
dnaJ	GCCATGAAA	14260	Chaperone
ftsK	GCTTCCATC	937143	Cell division
	CTGGCAGCC	934308	
hflC	GTATTGAAG	4403912	Protease regulator
miaA	ACGCTCTTC	4388831	tRNA modification
mnmE	GGGCTAAGT	3887200	tRNA modification
	GGCCGGGAA	3887126	
mnmG	CCCATGATG	3924896	tRNA modification
	AATCTGACC	3924521	
Mu *6*	GGTAAAAGG	4732	Unknown
Mu *kil*	GTCTTAATG	4508	Host cell death
Mu *9*	GCATTGTAT	5759	Unknown
recA	GGTGAGAAG	2822884	Recombination
	GGCTCATCA	2823258	
recC	TCCTGGCAC	2959669	Recombination
	GCTTTGACC	2960353	
	GGCTATGGC	2961293	
recG	GGCTTATGG	3825559	Recombination
	GTGTAGCTC	3826850	
ssrA	GCTTAGAGC	2755729	tmRNA
trmH	GCTGGGTAC	3824728	tRNA modification; polar effect on RecG
uvrD	GATCCACGC	3998683	DNA helicase
	CGTCTTACC	3998092	

[a] The 9-bp repeat was determined using a single primer that anneals within the transposon. Only one of the two junctions was sequenced, and we inferred the other junction because this transposon creates 9-bp repeats. The last nucleotide of the inferred repeat shows the last intact nucleotide of the gene, reading in the 5′ to 3′ direction with respect to the reference strand of the genome (regardless of the direction of the gene).

[b] Positions are given with reference to MG1655 sequence accession number U00096.3, with the exception of the Mu genes, which are given with reference to the sequence of bacteriophage Mu accession number AF083977.1.

identification of the exact location of some of the insertions is unique to this dissertation). Insertion mutants in genes *recA*, *recC*, *recG*, *ftsK*, *trmH*, *uvrD*, and three different bacteriophage Mu genes were the most hypersensitive, whereas mutants in *dinG*, *miaA*, *mnmE*, and *mnmG* showed more modest levels of sensitivity (for sensitivity data on *ssrA*, *dnaJ* and *hflC* mutants, see (84)).

One *trmH* knockout mutant was identified in the aza-C sensitivity screen. TrmH is a tRNA methyltransferase, and one could speculate that TrmH is thereby important in the activity of tmRNA. Alternatively, insertion mutants in *trmH* have been shown to have polar effects on *recG* (101). Because we identified *recG* knockout mutants in the screen, the aza-C sensitivity of the *trmH* insertion could be due to a similar polar effect. To distinguish between these models, we transformed the *trmH::kan* cells with a plasmid from the Aska collection overexpressing RecG (102). The *trmH::kan* cells with pCA24N-RecG were no longer hypersensitivity to aza-C (Figure 10 A). Moreover, *trmH* cells that were made by P1 transducing from the Keio collection, which creates an in-frame deletion of *trmH*, were not hypersensitive to aza-C (Figure 10B). We conclude that the *trmH* insertion is hypersensitive due to a polar effect on RecG expression.

Three insertions that led to strong aza-C hypersensitivity were localized within bacteriophage Mu genes (*kil*, *6* and *9*). The HK21 parental strain is not a traditional Mu lysogen, but it does have a *dinD*::*lacZ* fusion construct that is known to contain Mu

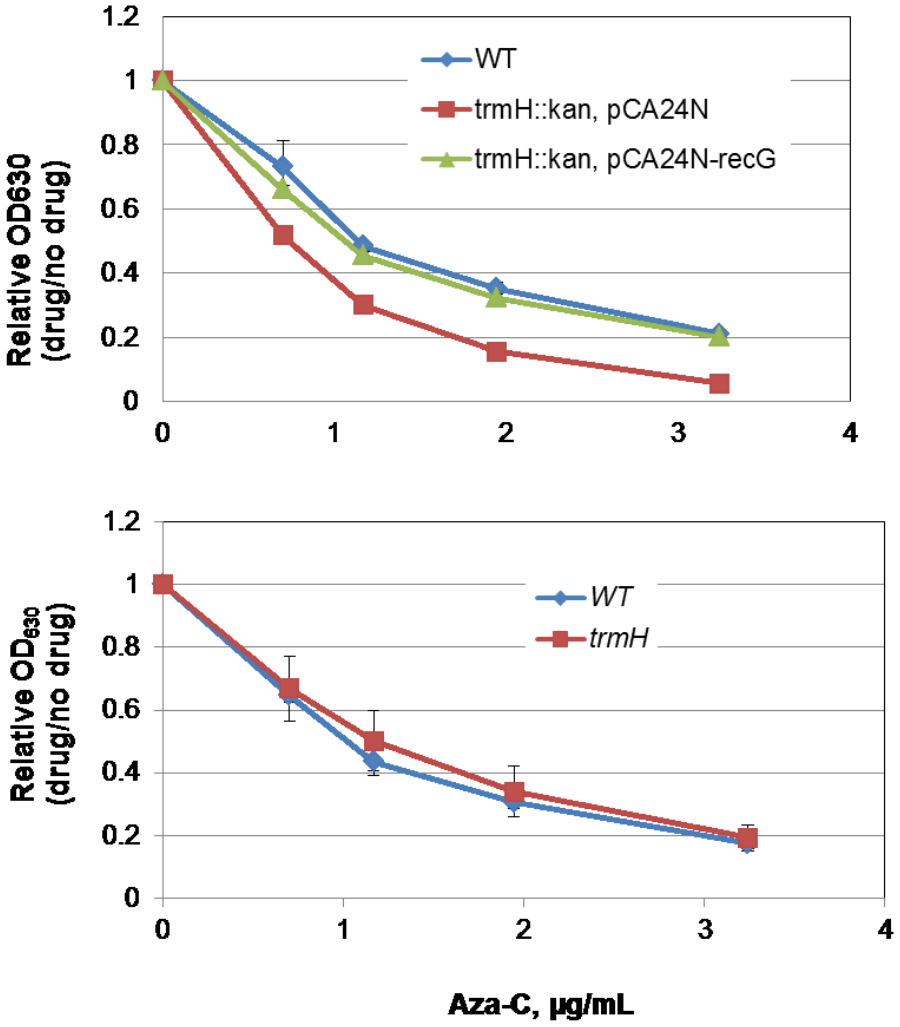


Figure 10: trmH::kan insertion is hypersensitive due to a polar effect on RecG expression

In Panel A, aza-C hypersensitivity profiles were calculated for *trmH::kan* mutants complemented with a plasmid overexpressing RecG or a vector control. RecG activity relieved the hypersensitivity of the *trmH::kan* insertion. In Panel B, *trmH* mutants were created using the Keio collection, creating an in-frame deletion of *trmH*. Error bars for each panel are the standard deviation for three independent cultures.

prophage DNA (103,104). For other reasons, we had subjected strain JH39, the *E. coli* strain from which the *dinD*::*lacZ* fusion was moved, to next-generation sequencing analysis. This confirmed the presence of phage Mu DNA, and allowed us to assemble the complete sequence of the *dinD*::*lacZ* fusion construct (see below) (Table 3). The three transposon insertion mutants were in Mu genes *kil*, *6* and *9,* with the transposon always in the orientation that aligns the kanamycin-resistance gene with the Mu early reading frames in this region. Each of the three insertion mutants could therefore presumably activate expression of Mu gene *gam*, which is only 42 bp downstream of the most distal transposon (the one in gene *9*), by read-through transcription. Gam is normally transcribed as part of an early operon during Mu infection, starting at the P_e promoter (105,106), but would not normally be expressed in lysogens (or presumably from the *dinD*::*lacZ* construct) (107). The Mu Gam protein inhibits nucleases, including RecBC, by binding to DNA ends (108). We therefore infer that these three insertions activate Gam expression, essentially creating a phenocopy of a RecBC knockout mutant (Shee *et al.* (109) recently showed that Gam expression induces *recBC* knockout phenotypes). Consistent with this model, the HK22 Mu *9*::*kan* cells are hypersensitive to both nalidixic acid and ciprofloxacin, well-documented phenotypes of RecBC knockout mutants (Figure 11). Expression of the toxic product of the Mu *kil* gene should not occur in any of the three mutants because the insertions are within or downstream of *kil*.

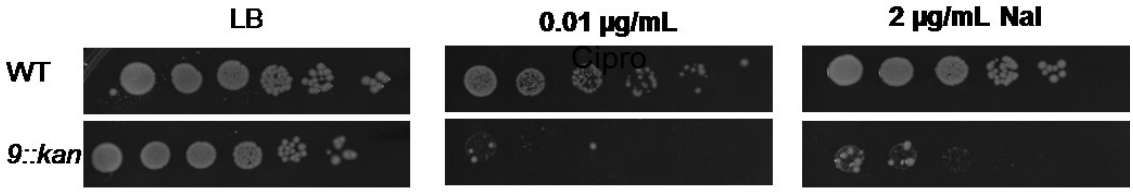


Figure 11: Spot test for quinolone hypersensitivity of strain with insertion in Mu *9*
Wild-type or the insertion mutant with a transposon in Mu gene *9*, both containing pBAD-MEcoRII, were diluted from overnight culture to approximately 4 x 10^8 cells/ml. Ten-fold serial dilutions were generated across a 96-well plate and 5 µl of each dilution was spotted onto LB plates with no drug (left panel), ciprofloxacin (Cipro, 0.01 µg/mL, middle panel) or nalidixic acid (Nal, 2 µg/mL, right panel). Plates were photographed after overnight incubation at 37°C.

dinD::lacZ fusion sequence

The sequence of the *dinD*::*lacZ* fusion was determined by deep sequencing of DNA from strain JH39, followed by reconstructing the genome sequence using the alignment program Geneious and NCBI Blast. The sequence of the fusion is summarized in Table 3, and can be reconstructed by accessing the sequences and positions listed in descending order in the Table. The order of the numbers in the "Positions" column represents the direction of each sequence fragment in the fusion.

The *dinD*::*lacZ* fusion was originally created by infecting cells with the Mud(Ap^R, *lac*) phage and isolating strains that induced β-galactosidase activity during mitomycin C treatment (104,110). The Mud(Ap^R,*lac*) phage contains a temperature-sensitive mutation in the Mu repressor (*cts62*) which induces phage gene transcription at high temperatures, making the cells temperature sensitive for growth. Strain JH39 was created using the *dinD1*::MudI(Ap^R*lac*) fusion, but with selection for derivatives that had lost the temperature sensitivity and the ability of the Mu phage to transpose (104,111). The sequence shows an IS1 insertion in gene *A* of Mu, which explains the loss of transposition. This IS1 element has a 298-bp deletion, with an area of microhomology near the deletion site that may have been involved in the deletion formation. The Mu sequence still contains the *cts62* mutation, and therefore another mutation(s) must suppress the temperature sensitivity. Likely candidates are: (1) the IS1 insertion in gene *A*, because IS1 (including this deletion derivative) carries a transcriptional terminator

Table 3: Sequence summary of the *dinD::lacZ* fusion construct

Organism	Accession #	Features	Mutations/Deviations	Position
MG1655	U00096.3	*dinD*		3817760 - 3818210
Bacteriophage Mu	AF083977.1			36717 - 36670
Bacteriophage Mu	AF083977.1			36662 - 36705
MG1655	U00096.3	Part of *trpA*, all of *trpB*	Possible C insertion at 1316460, 1316454, A/T ambiguity at 1316456	1316464 - 1315071
MG1655	U00096.3	*lacZ*, *lacY*, *lacA, cynX*, *cynS*		366297 - 359483
Unknown			Insertion: TTCGCGCGCC	
E. coli Tn3	V00613.1	Ampicillin resistance		3571-4957
Bacteriophage Mu	AF083977.1		C19284G[a]	25884 - 16130
E. coli K12 mobile genetic element IS30[b]	X00792.1			1221-1
Bacteriophage Mu	AF083977.1	Includes *gam* and gene *9*		16131 - 3183
E. coli insertion sequence 1[c]	V00609.1	IS1 variant		768-758
E. coli insertion sequence 1[d]	V00609.1	IS1 variant	A/G ambiguity at 127, G396A, T393G, A304G, G301A, G298T, G262T	460-1
Bacteriophage Mu	AF083977.1			3190 - 1
MG1655	U00096.3	*dinD*		3818206 - 3818584

(112) and the toxic Mu gene *kil* (also called gene *5*) is downstream; and (2) the IS30 element, which interrupts the late gene region of Mu and might thereby prevent expression of one or more toxic late genes. The Mu sequence in this element reveals an inverted repeat of 43 bp near the C end of Mu, most likely due to non-standard fusion formation during the transposition of the MudI(Ap^R*lac*) element into *dinD* or the fusion/deletion of Mu and *trpB* (the latter mutation is called ΔFI by Kenyon and Walker; non-standard fusion formation is described in (113)). The Mu sequence also contains a point mutation in gene *T*.

Candidate approach: Functions potentially involved in DPC repair

Because the screen proved to be non-saturating (i.e- *smpB* mutants are clearly hypersensitive to aza-C treatment yet were not identified in the screen), we also conducted a candidate-based screen to directly address the involvement of proteins that might be expected to impact survival after DPC formation. Deletion mutations in a variety of genes were moved, usually from the Keio knockout collection (114), into the HK22 background by phage P1-mediated transduction. Transductants were verified by PCR and the pBAD-MEcoRII plasmid was transformed into the strain in order to test aza-C sensitivity.

One set of genes tested were those whose products might be expected to be directly involved in the repair or tolerance of DPCs or downstream DNA damage. The

SbcCD complex has been shown to remove proteins bound to DNA *in vitro* by introducing double strand breaks (115), and one could imagine a repair pathway similar to that used to trigger meiotic recombination near SPO11 complexes (116). UmuCD (also known as pol V) is a translesion polymerase that is expressed during the late phases of the SOS response (117), and perhaps base insertion opposite the DPC could be involved in some tolerance pathway. RecF is part of the RecFOR recombination pathway, which is alternative to the RecBC pathway for RecA-mediated recombination (118). RecFOR is important in stalled fork processing after UV damage (119,120), and aza-C-induced DPCs cause fork stalling (42). Knockout mutants for each of these functions expressing M.EcoRII displayed essentially wild-type sensitivity to aza-C in cells in the microtiter plate assay (Figure 12). We conclude that none of these functions play important roles, or if they do, there are overlapping alternative functions that can take their place efficiently.

Both *dnaJ* and *hflC* knockouts are hypersensitive, raising the intriguing possibility that these proteins could be involved in proteolysis of the protein that is covalently linked to DNA (84). Since DnaJ and DnaK frequently act together (121,122), we also generated a *dnaK* knockout mutant and tested its sensitivity to aza-C. We were surprised to find that the *dnaK* mutant is somewhat resistant to aza-C treatment (Figures 8, 12 and 13). This implies that the *dnaJ* mutant sensitivity is due to a DnaJ activity that is DnaK-

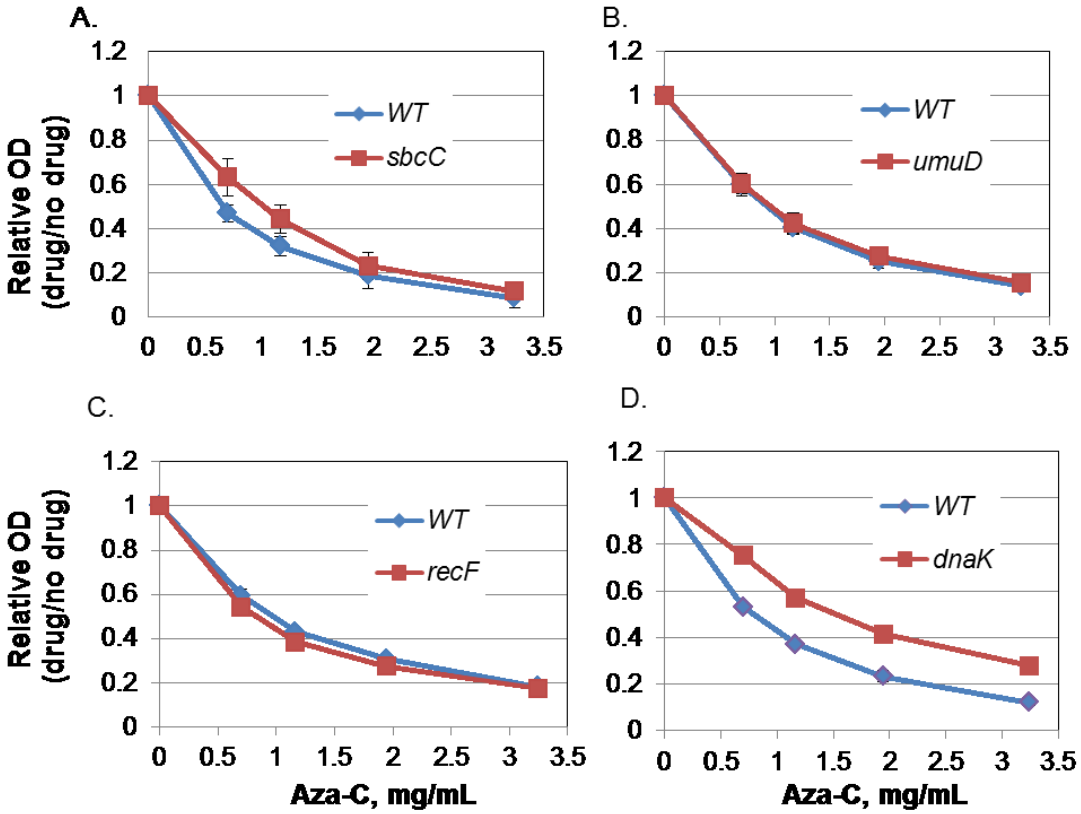


Figure 12: Aza-C hypersensitivity of potential repair proteins.
The indicated cells were grown and drug titration curves were calculated as in Figure 2, using an endpoint time of 50% of the max growth rate. For each experiment, wild-type and the indicated mutant was grown in three independent cultures with the exception of *sbcC* mutants, where 6 independent cultures were grown. Error bars are the standard deviation for the respective number of independent cultures.

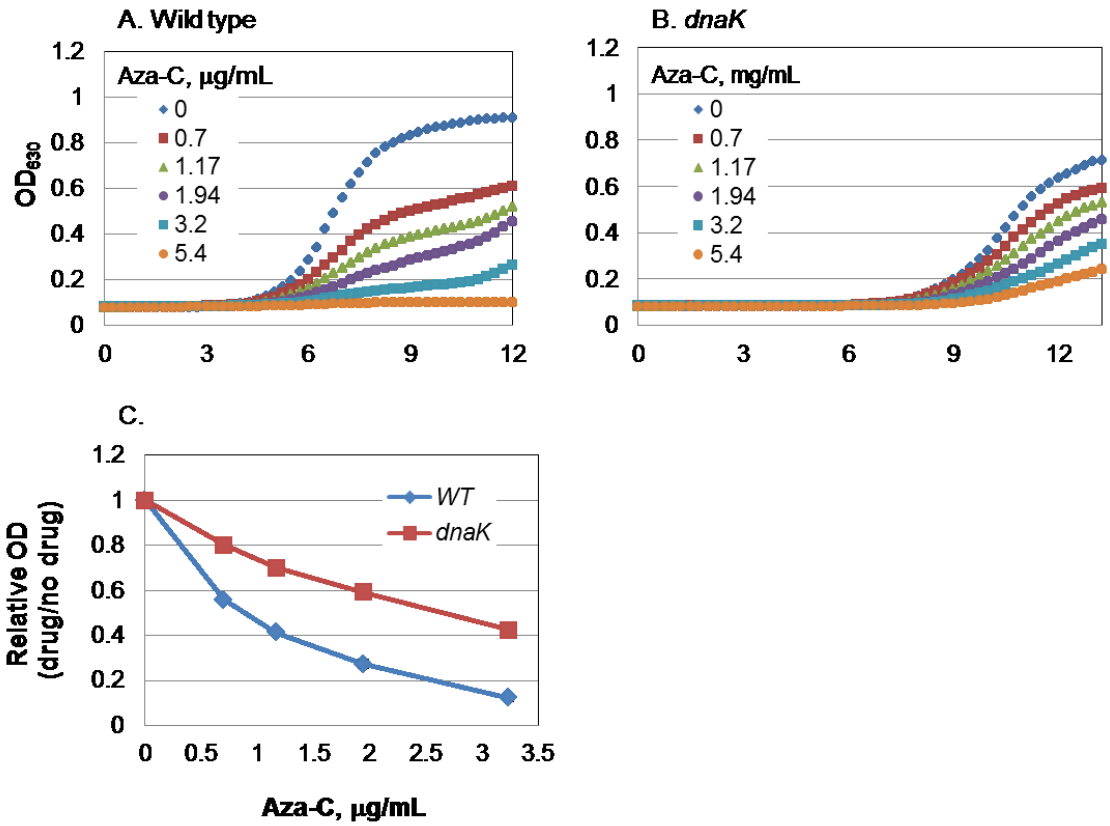


Figure 13: Aza-C hypersensitivity profiles for *dnaK* cells
Panels A and B are growth curves for WT and the slow growing *dnaK* cells, respectively. The growth curves in Panel B were used to calculate the drug titration curve in Figure 12. Panel C is the drug titration curve using the same *dnaK* cells and calculated similarly to the curves in Figure 12, with the exception that OD560 were plotted from the time at which the growth rate of the no-drug culture dropped to 90% of the earlier exponential rate.

independent, and perhaps DnaK normally sequesters DnaJ and reduces its ability to function in a DPC-relevant pathway (also see Discussion).

2.3 Discussion

The data presented here expands upon our previously proposed "chain-reaction" model during DPC formation in which RNA polymerase is blocked by the DPC, and then translating ribosomes on the attached nascent transcript are blocked by the stalled polymerase and potentially backed up in an array of stalled ribosomes behind the RNA polymerase. We have shown here that the tmRNA system functions to aid in cell survival by clearing the stalled ribosomes. Another question that we address is how tmRNA is able to access the empty A-site on the ribosome for the trans-translation process. Presumably, either the RNA at the A-site is cleaved to release the ribosome-RNA (5′ end) complex from the stalled RNA polymerase-RNA (3′ end) complex, or the RNA polymerase and nascent RNA are released first, allowing translation to proceed to the premature end of the transcript. A-site cleavage normally occurs by a two-step pathway when ribosomes are stalled in the middle of a transcript, such as at rare codons (89,90,123,124). First, RNase II 3′-5′ exonuclease activity is required to degrade the message to the leading edge of the ribosome (although the 3′ end would be blocked in the case of a DPC-stalled transcription complex). This degradation then facilitates a second nuclease activity, presumably by the ribosome

itself. We found that knockout of RNase II, RNase D, RNase PH, or RNase LS did not cause hypersensitivity to aza-C, arguing against the A-site cleavage model for tmRNA function after DPC formation.

The second model proposes that RNA polymerase and nascent transcript are released by some active process, to license tmRNA action on the released transcript. Presumably, release and reutilization of RNA polymerase would be important even if tmRNA gains access to the transcript through A-site cleavage, so the fate of the RNA polymerase is relevant in any case. *In vitro* studies have shown that RNA polymerase is blocked by aza-C-induced DPCs (43), and also that RNA polymerase blocked by a tightly bound non-cleaving mutant EcoRI endonuclease is stable for up to 1 hour (75). These results imply that release of RNA polymerase requires some additional release factor(s) whose identity remains obscure.

A well-studied mechanism for releasing RNA polymerase complexes blocked by certain obstructions in the template DNA involves the Mfd helicase (66,125). However, our lab previously found *mfd* mutants were not hypersensitive to aza-C, suggesting that Mfd is not involved in releasing transcription complexes blocked by DPCs (84).

We also tested for the possible involvement of transcription termination factor Rho by asking whether the Rho inhibitor bicyclomycin acts in a synergistic fashion with aza-C. We found no evidence for any drug interaction, arguing against Rho involvement. Interestingly, Rho induces termination and release of RNA polymerase

blocked *in vitro* by a tightly bound protein (75). In that case, however, the nascent RNA was not bound by ribosomes, which are known to inhibit Rho action (3). In summary, neither Mfd nor Rho appears to play an important role in surviving transcriptional blockage by aza-C induced DPCs.

In the aza-C hypersensitivity screen conducted previously, we found that mutational inactivation of helicase DinG or UvrD causes hypersensitivity to aza-C-induced DPCs, while inactivation of Rep does not (Figures 2, 4, and 5). While *dinG* and *uvrD* mutants are hypersensitive, the amount of DPC-induced SsrA tagging was not reduced in these mutants. Therefore, DinG and UvrD do not appear to be necessary to release the nascent transcript to allow tmRNA action (unless they have overlapping roles, but note that single mutants are hypersensitive). It remains possible that DinG and/or UvrD releases stalled RNA polymerase at a DPC, but that this step is not needed for tmRNA action.

HepA (RapA), a member of the SWI/SNF superfamily of helicase-like proteins, is an RNA polymerase binding protein that allows recycling of the polymerase at poorly characterized post-transcription/post-termination complexes that form *in vitro* (76,126). Given its recycling function, HepA presented as a candidate for releasing RNA polymerase and the nascent transcript from the DPC-blocked complexes. However, a HepA knockout strain was found to be no more sensitive to DPC formation than wild

type cells (Figure 2), even with concomitant inhibition of Rho activity (Figure 3), arguing against an involvement of HepA.

E. coli RNAP can transcribe through reversibly bound proteins *in vitro (82,99)*, and T7 RNAP can transcribe through a DPC *in vitro*, albeit with low frequency and in a mutagenic fashion (127). We were therefore interested in the possible role of factors that might promote polymerase read-through at blocking lesions. Transcription factors GreA, GreB and DksA interact with the secondary channel of RNA polymerase, and have been shown to decrease pausing and to stimulate elongation past a lesion or protein roadblock (80,82,128). We found that *greA* and *greB* mutants are no more sensitive to DPC formation than wild type cells, even when termination is decreased by decreasing Rho activity (Figures 2 and 3).

The one transcription factor that we did find to be somewhat protective against aza-C-induced DPCs is DksA, but only when measured late in growth (Figures 2, 4, and 6). DksA modulates transcription elongation outside of its transcription initiation roles, although the specific mechanism is not well defined. Perhaps most relevant to the present study, it has been proposed that DksA prevents transcription-replication conflicts by destabilizing paused transcriptional elongation complexes (73,81). This factor does not appear to be important in releasing nascent transcripts from the DPC-stalled complex for tmRNA action, because the knockout strain had increased rather than decreased SsrA tagging (Figure 6). The increased tagging observed in the DksA

knockout mutant suggests that DksA activity might instead be in competition with the factor(s) responsible for creating the substrates for the tmRNA system. One intriguing possibility is that DksA promotes transcription through DPCs, allowing bypass of the lesion when it is located on the template and/or the non-template strand.

We have also considered the possibility that RNAP release and DPC repair are somewhat coupled, and have attempted to identify the repair pathway(s) for DPCs. This work has expanded on the aza-C hypersensitivity screen by identifying the exact positions of each of the transposon insertion sites, including those in the Mu transposon. In order to best understand the placement of Mu DNA in HK22, as well as to have a better understanding of the commonly used *dinD::lacZ* fusion, we also compiled the *dinD:lacZ* fusion sequence in its entirety.

In addition to the aza-C hypersensitivity screen we sought to test other candidate mutants that might be expected to impact survival after DPC formation. Several of the genes identified from the transposon screen were previously implicated in genome stability, including the genes that encode recombination proteins RecA, RecC and RecG (Table 1). These proteins could play a direct role in resolving DPCs, or alternatively, they could be involved in repairing DNA ends that are secondary effects of DPCs (e.g. breaks generated at stalled replication forks; see (42,129,130)). Even though aza-C-induced DPCs lead to fork stalling, we detected little or no effect on aza-C sensitivity from

knocking out the RecF protein, which plays a key role in the consequences of fork stalling from UV-induced lesions (119,120).

On the other hand, recombination proteins could potentially play a more direct role in DPC repair. An attractive model is provided by the pathway of meiotic recombination, where Mre11/Rad50 complex cleaves DNA near an SPO11-DNA covalent complex to initiate a double-strand break repair event (116). This pathway was also implicated in repair of topoisomerase cleavage complexes (116). The homolog of the Mre11/Rad50 complex in *E. coli* is SbcCD, although the role of this protein is quite distinct from that of Mre11/Rad50 (131,132). Strikingly, SbcCD can cleave DNA *in vitro* near proteins that are very tightly bound to the end (115), and so we asked whether SbcCD might be involved in repair of DPCs. However, we found no change in aza-C sensitivity in a SbcC knockout mutant, strongly arguing against this possibility.

DnaJ functions as a co-chaperone to the *E. coli* Hsp70 protein DnaK but also has DnaK-independent activities (121,122). To further explore the role of DnaJ in aza-C sensitivity, we also generated and tested a *dnaK* knockout mutant. The *dnaK* mutant turned out to be modestly resistant to aza-C treatment (Figure 3D and Supplemental Figure 4A). This indicates that the role of DnaJ in aza-C sensitivity is DnaK-independent, and perhaps also that DnaK can sequester DnaJ and reduce its ability to function in a DPC-relevant pathway. A different interpretation is that *dnaK* mutants overproduce heat shock proteins (133), and perhaps one or more heat shock proteins assist in survival after

aza-C. In general, given the broad roles of HflC, DnaJ and DnaK in protein metabolism, it will be important to decipher whether aza-C sensitivity is modulated by direct or indirect effects.

In summary, our studies have identified DksA as a potential factor that assists transcription elongation during DPC formation, presumably by aiding in transcription past a lesion on the non-template strand. The factors DinG and UvrD are also important factors for protection against DPCs, and these proteins might be involved in releasing stalled RNA polymerase. We provided evidence against involvement of several RNases, Rep, Mfd, Rho, GreAB and HepA in any pathway that is protective against aza-C-induced DPCs. We have yet to discover pathway by which the tmRNA system gains access to ribosomes within DPC-stalled transcription/translation complexes, further emphasizing the knowledge gap that exisits for tmRNA biology.

2.4 Materials and Methods

Materials

Aza-C was obtained from Sigma-Aldrich, nitrocellulose membranes (Protran® BA 85) from Whatman, and polyclonal M.EcoRII antibody from Proteintech Group, Inc (custom generated). LB broth contained Bacto tryptone (10 g/L), yeast extract (5 g/L) and sodium chloride (10 g/L), supplemented with the appropriate antibiotics. Antibodies against the SsrA tag was kindly provided by Tania Baker (MIT).

***E. coli* strains**

HK21 is a derivative of strain ER1793 (obtained from New England Biolabs) and has the following genotype: F$^-$ *fhuA2 Δ(lacZ)r1 glnV44 e14$^-$(McrA$^-$) trp-31 his-1 rpsL104 xyl-7 mtl-2 metB1 Δ(mcrC-mrr)114::IS10) ΔsulA (Keio deletion) dinD::lacZ.* Western blots were done using the HK22 strain (84). *E. coli* strains for the candidate gene approaches were constructed by P1 transduction of deletions from the Keio collection (114) into strain HK22 (same as HK21 except lacking the *dinD::lacZ* fusion), selecting for the kanamycin resistance marker inserted at the Keio deletion site. Plasmid pBAD-MEcoRII was then transformed into each strain following confirmation of the gene knockout.

Plasmids

Plasmid pBAD-MEcoRII was used to express M.EcoRII(84). This plasmid is a pBAD33 derivative (98) which carries chloramphenicol resistance and the M.EcoRII coding sequence under control of an arabinose-inducible promoter. Expression of M.EcoRII from this plasmid is repressed with glucose (0.2%) and activated with arabinose (0.05%). Plasmid pR234 is a pKK223-3 derived plasmid with the M.EcoRII gene controlled by a P_{tac} promoter (87). Plasmid pRK1, a vector control for pR234, was constructed by cleaving pR234 with BamH1 and religating, which removes the entire M.EcoRII coding sequence. pCA24N-recG is from the ASKA collection of E. coli ORF without the GFP or His tag (102). RecG is expressed from the $P_{T5\text{-}lac}$ promoter that can be activated by IPTG.

Growth kinetics for aza-C sensitivity

Aza-C sensitivity was measured with continuous growth curves in a temperature-controlled ELx808™ Absorbance Microplate Reader. Overnight cultures in LB media at 37°C were diluted to an OD_{630} of 0.5, diluted 1:2000 in LB with chloramphenicol and 0.05% arabinose, and then mixed with an equal volume (75 μl each) of LB containing serial dilutions of aza-C (15 μg/ml) in 96-well plates. The cells were grown in the plate reader at 37°C with constant shaking, and OD_{630} was measured every 15 minutes for 18 hours.

Western blots for SsrA tagging

SsrA tagging levels were analyzed in mutant derivatives of the HK22 pBAD-M.EcoRII strain background. Cells were pregrown overnight in LB media at 37°C in the presence of 0.2% glucose, and then diluted to an OD_{560} of 0.1. The cells were then grown to an OD_{560} of 0.5 in the same media, with or without aza-C (50 μg/ml). Cells were then harvested by centrifugation, washed once with LB, and finally resuspended in LB containing arabinose (0.05%) to induce M.EcoRII expression. After one hour at 37°C with shaking, cell samples equivalent to 2 ml of OD_{560} = 0.5 were harvested, incubated in an ethanol/dry ice bath for at least 15 minutes, and then stored at -20°C.

Frozen cell pellets were thawed at room temperature and resuspended in 25 μl of water and 25 μl of sample buffer (20% glycerol, 100 mM Tris pH 6.8, 2% SDS, 2% β-mercaptoethanol, and bromophenol blue), and boiled for 5 min in a water bath. An

aliquot (15 µl) of each sample was loaded onto a 7.5% polyacrylamide (Tris-HCl) gel and run for approximately 2 h in 25 mM Tris-Glycine buffer containing SDS (0.1%). The portion of the gel containing proteins larger than about 75 kDa was cut off and stained with Coomassie blue dye to serve as a loading control. The remaining portion of the gel was transferred to a nitrocellulose membrane for 60 min at 12 V using a Genie Blotter transfer device (Idea Scientific Co.). The blot was blocked for 1 h in 20% non-fat milk powder solution (Biorad) in Tris-buffered saline (TBS). The membrane was incubated overnight at 4°C with polyclonal *ssrA* primary antibody and Tween (0.1%), and then washed three times with TBS buffer at room temperature (10 min each). The membrane was incubated with secondary antibody (IRDye 800CW-conjugated goat anti-rabbit IgG (LI-COR®)) for 1 hour, and the washes were repeated. After air-drying, the membrane was scanned on an Odyssey Imaging System (LI-COR Biosciences), and quantified using the Odyssey software.

Spot tests for aza-C and quinolone sensitivity

Overnight cell cultures in LB (plus chloramphenicol) were diluted to roughly 4 x 10^8 cell/ml. Ten-fold serial dilutions were then generated across a microtitre plate and 5 µl of each dilution was spotted onto LB plates with appropriate antibiotics and the indicated aza-C, ciprofloxacin, or nalidixic acid concentration. The plates were incubated overnight at 37°C.

3. Examining the mechanisms of elongation inhibition for Actinomycin D and Streptolydigin

The work for the Actinomycin D portion of this chapter was assisted by Ramsey Al-Khalil, an undergraduate researcher in our lab.

3.1 Introduction

In the previous Chapter, we found that the biology at the sites of DPCs is quite complex, with various factors competeing to assist with the stalled replication, transcription, and translation, as well as the repair of the DPC. Therefore, an alternative approach was necessary to specifically isolate the effect on transcription. For the purposes of these studies we chose Streptolydigin and Actinomycin D. However, the precise mechanisms of inhibition by these two inhibitors are not fully understood, and therefore in this Chapter we address some of the uncertainties that could complicate the study of transcription.

Streptolydigin is a transcription elongation inhibitor that inhibits RNA chain elongation (57-59). Stl binds directly to RNAP, interacting with both the β and β′ subunits to stabilize the straight-bridge-helix RNAP-active-center conformational state (60,61). This stabilization prevents normal conformational changes necessary for translocation of the growing nascent chain (60,61). Stl binding does not appear to alter nucleotide binding to RNAP, phosphodiester bond formation, or DNA translocation (61). Elongation complexes stalled by Stl are very stable *in vitro;* complexes are still bound to DNA after 50 minutes of high Streptolydigin treatment, and removal of Stl

results in resumption of transcription after a delay *in vitro* (59). However, the stability and duration of Stl-RNAP interactions *in vivo* is uncertain, as are the specific requirements for Stl binding to RNAP.

We are also interested in the fate of elongation complexes stalled by the eukaryotic transcription inhibitor Actinomycin D. ActD is a small molecule with antibiotic and anticancer capabilities (62,63). It is primarily known as a DNA intercalator, with the ability to intercalate into double stranded DNA, bind to single-stranded DNA, and hemi-intercalate into single stranded DNA. It is hypothesized that ActD binds to the distorted DNA at the ends of the transcription bubble, preventing RNAP elongation, but the exact mechanism of this inhibition is unknown (63). For instance, it remains to be answered whether ActD impedes RNAP translocation simply due to DNA intercalation, or if it directly interacts with RNAP as well as DNA.

In this Chapter, we address several of the questions surrounding elongation inhibition by Stl and ActD. These topics include the stability of RNAP-Stl interactions, the effect of pausing frequency on the effectiveness of the inhibitors, and the potential for ActD-RNAP interactions. We will show that Stl binding is transient, and that pausing frequency directly correlates with both Stl and ActD efficiency. We isolate several RNAP mutants that are resistant to ActD treatment and propose that the method of RNAP inhibition by ActD is more than just simple steric hindrance due to DNA intercalation.

3.2 Results

Sensitizing *E. coli* cells to Streptolydigin and Actinomycin D

E. coli cells are resistant to both Actinomycin D and Streptolydigin and require modifications to the cell membrane to become sensitive. TolC is an outer membrane porin that is the outer membrane component of several multi-drug efflux systems; *tolC* mutations in *E. coli* cause susceptibility to a variety of antibiotics, including Streptolydigin (60). For our Stl studies we primarily used EW1b cells, which are known to have a disruption in the *tolC* gene. When EW1b cells were first isolated it was thought that there was a deletion in the *tolC* gene , as no revertants were found by the author (134). However, PCR of the *tolC* gene in EW1b cells shows a 1 kb increase in size compared to the wildtype *tolC* sequence, leading us to believe the disruption was an insertion. In order to identify this insertion, we sequenced the full *tolC* gene in EW1b and found that there is an IS5 insertion 191 bp into the *tolC* gene in EW1b cells.

To make *E. coli* sensitive to ActD, we first attempted to treat cells with EDTA in a checkerboard assay similar to the bicyclomyin/aza-C double-drug assay used in Chapter 2, as sub-inhibitory levels of EDTA increases susceptibility to ActD (135). We were successful in making *E.coli* sensitive to ActD; however, cell growth was inconsistent using this method and results were thus hard to interpret. To establish a more controlled assay system, we began testing gene knockouts that could also modify the cell susceptibility to antibiotics. We first tested MG1655 *tolC* cells for ActD

sensitivity; however, these cells proved to be as resistant as MG1655 cells. Communication with Dr. Meta Kuehn led us to test *pal* and *mrcB* mutants, as these had been shown in her lab to be sensitive to ActD treatment. Pal is a peptidoglycan-associated lipoprotein that is a part of the Tol/Pal system, which functions to maintain cell envelope integrity. MrcB is an inner membrane enzyme catalyzing the transglycosylation and transpeptidation of peptidoglycan precursors; accordingly, MrcB mutants have decreased peptidoglycan layer density leading to hypersensitivity to certain antibiotics. Growth sensitivity tests with MG1655 *pal* and MG1655 *mrcB* cells confirmed that these cells are sensitive to ActD (Figure 14). MG1655 *mrcB* cells grew more consistently in our ActD sensitivity tests conducted in the microtiter plates across several assays on different days, and MG1655 *pal* cells were four-fold more sensitive in solid plate assays (where much greater volumes of ActD are needed). Therefore both cells lines were utilized where technically advantageous.

Stl does not cause replication-transcription collisions, consistent with transient pausing of RNAP

It has been shown that washing out Stl in *in vitro* studies results in resumption of elongation, implying that Stl only transiently binds to RNAP (59). We wanted to confirm that this is true for *in vivo* RNAP-Stl interactions. To do so, we employed the method of Tehranchi *et al.* in using SOS induction as an indicator for transcription arrest.

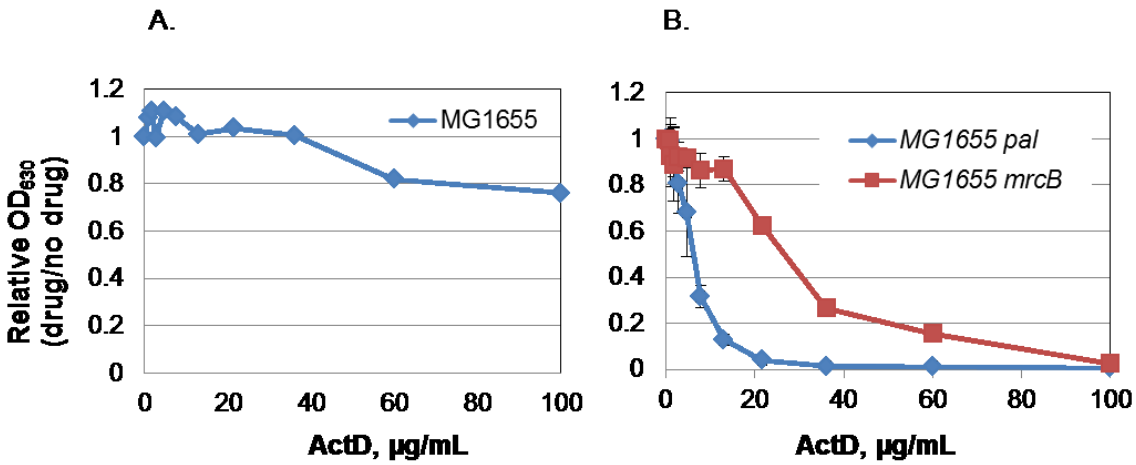


Figure 14: MG1655 *pal* and *mrcB* mutants are sensitive to Actinomycin D
Panel A is a representative drug response curve for MG1655 cells, which are not sensitive to ActD treatment. In Panel B, MG1655 *pal* and *mrcB* mutants were tested for ActD sensitivity using the same method. Deletions of both *pal* and *mrcB* make the membrane more permeable and allow ActD to penetrate the cells. Error bars are the standard deviation for three independent cultures.

They used serine hydroxamate (SHX), which causes amino acid starvation and subsequent translation and transcription arrest, to examine the effect of DksA activity on transcription and replication. They found that removing DksA during SHX treatment led to replication-transcription collisions. These collisions resulted in DNA damage that induced the SOS response (81). If Stl is forming a permanent or stable interaction with RNAP, it should also result in replication-transcription collisions and downstream SOS induction. However, if the pausing is only transient, there is most likely enough time for replication to resume before fork collapse. In this scenario there would be no induction of the SOS response.

To test for SOS induction during Stl treatment we performed quantitative RecA western blots on cells treated with Stl. As shown in Figure 15, we found no increase in RecA levels during Stl treatment (Mfd results are addressed in Chapter 4). This strongly suggests that Stl binding is transient *in vivo.*

In vitro elongation reactions stalled by Stl and ActD

The effect of adding Stl and ActD to elongation complexes i*n vitro* have been studied before, yet these systems always employed total genomic DNA as a substrate. We wanted to monitor the effect of these drugs on transcription in cleaner system using a carefully controlled PCR product as a DNA template. We used a PCR fragment from the pDE13 plasmid which contains a transcript with a λ_{PR} promoter and t_{R2} terminator

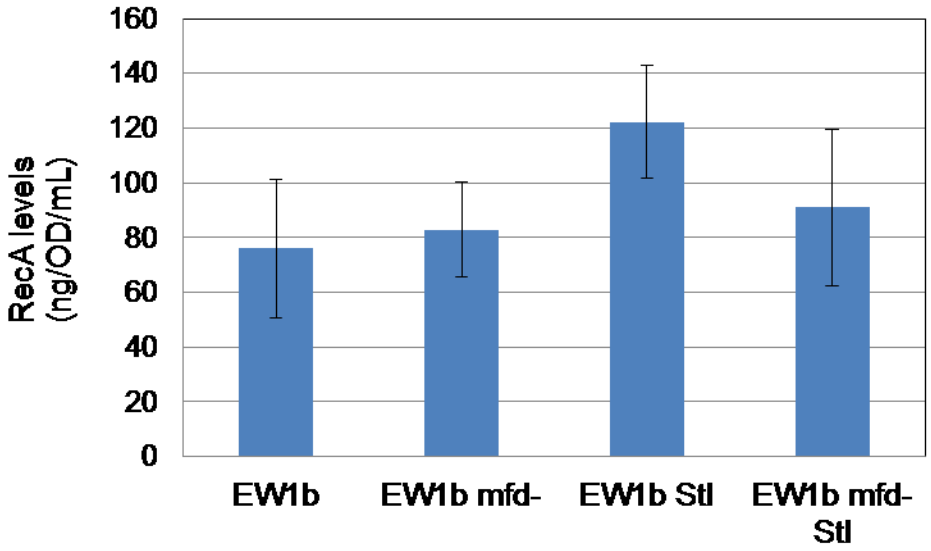


Figure 15: Stl treatment does not induce the SOS response
EW1b or EW1b *mfd* cells were grown in LB with or without Stl treatment as described in Materials and Methods. A standard curve of RecA protein was run on the same gel with the samples for accurate quantitation. RecA amounts for each sample were calculated by equating band intensity with a RecA amount, then adjusting for starting cell concentration and sample dilution (calculated from the BCA quantitation) to give RecA levels/OD/mL. Three repetitions for each sample were measured, and the error bars reflect the standard deviation of these samples.

(136). Transcription from this template in the absence of CTP allows for a stalled elongation complex at +24. The general protocol was to make open complexes by mixing DNA and RNAP in transcription buffer and preincubating at 37 °C. Then heparin and ATP, CTP and radioactive GTP in buffer were added at low levels to have one RNAP stalled at +24 per DNA molecule. All 4 nucleotides were added at saturating levels to allow for full length product formation from complexes not inhibited by the respective drugs.

Surprisingly, the addition of Stl did not lead to any inhibition of RNAP (Figure 16A), even at levels of Stl that were able to completely block transcription *in vitro* using coliphage T4 DNA as a substrate (59). Several attempts at troubleshooting the system were made, such as omitting Mfd buffer (which has fairly high glycerol content), adding more Stl, room temperature vs 37 °C, using less ATP, and omitting heparin. Under every condition tested Stl failed to inhibit RNAP. We propose that Stl is unable to bind to rapidly elongating RNAP, and instead must wait for RNAP to pause or otherwise be held in the "open" conformation due to structural considerations (See Discussion).

We next attempted similar tests *in vitro* with Actinomycin D. Unlike Stl, ActD efficiently inhibited elongation in our system. As shown in Figure 16B, ActD inhibited nearly all full length transcript production at 0.8 μM. Instead, a new band appears at roughly +50, indicating strong ActD inhibition of transcription at this location. In the

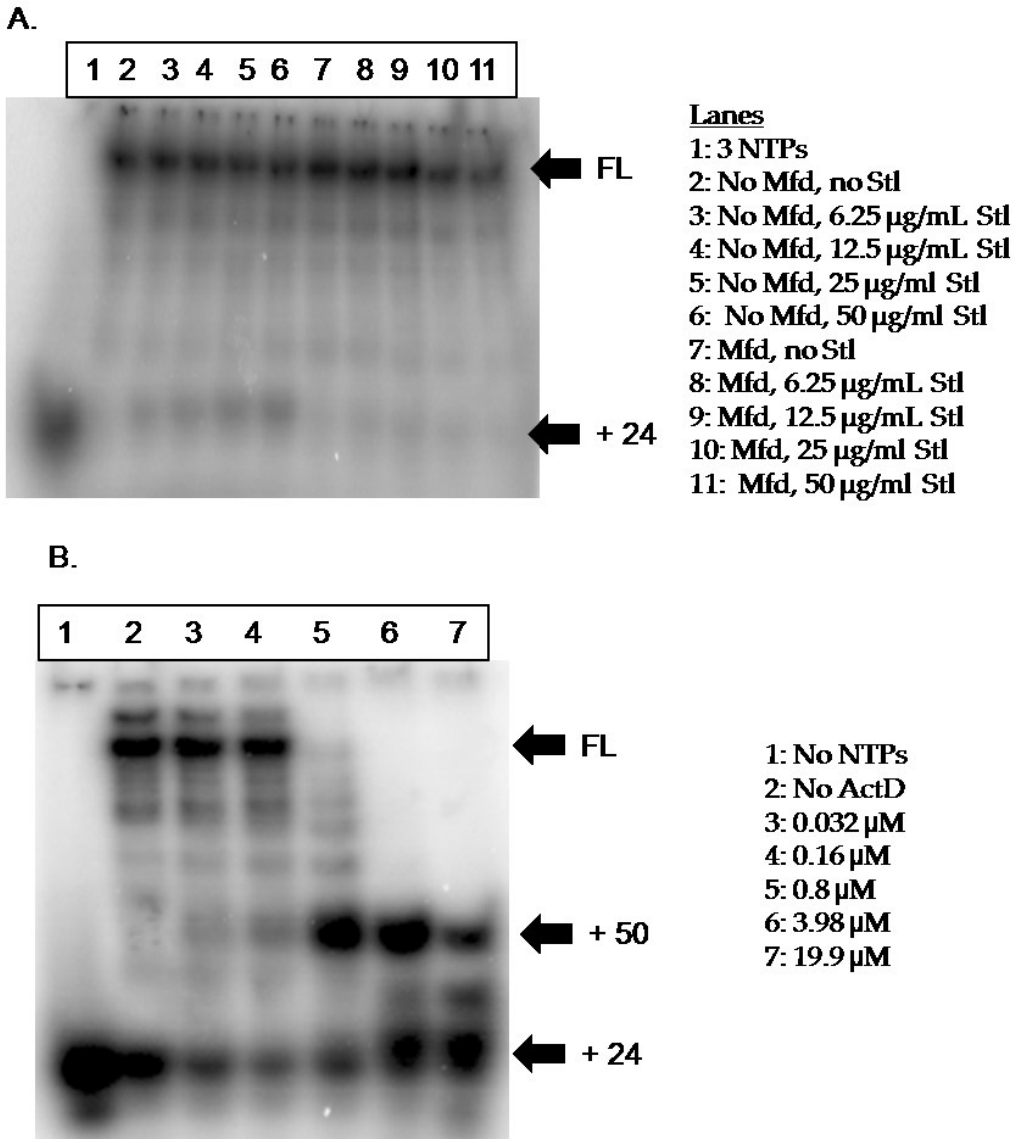


Figure 16: In vitro elongation assays with inhibitor treatment

Panel A is an *in vitro* assay monitoring the effect of Stl and Mfd treatment. Transcription was carried out using radioactive GTP, which allows for visualization of resulting RNAP products. FL= Full length product, + 24= Complex stalled or released at +24 on th transcript. Number list on the side of the gel indicates the amount of Stl added to each sample. 400 nM Mfd was added where indicated. In Panel B, *in vitro* reactions testing the ability of ActD to stall elongation complexes were conducted. FL= Full length product, +50= Complex stalled or released at +50, + 24= Complex stalled or released at +24. Numbered list on the side of the gel indicates the amount of ActD added to each sample.

pDE13 template there is a GCGC at position +53, which would be expected to be a strong ActD binding site due to increased hydrogen bonding (64) and would explain the band at +50. Other GC pairs exist in the template which could lead to the faint bands seen in the lanes.

We were surprised to find that there is a very tight window for the titration of ActD. In Figure 16B, full length production was essentially unaffected at 0.16 μM (the next lowest level tested), and additional assays have shown that full length transcript levels at 0.67 μM are roughly the same as untreated reactions. It has been shown that ActD has very slow on- and off- rates of binding, and the tight titration window is consistent with the idea that once ActD binds to DNA its inhibition of RNAP is very stable.

Modulating RNAP speed affects cell sensitivity to Stl

If Stl binding requires RNAP to be paused or otherwise held in an open conformation, it stands to reason that altering the pausing frequency of polymerase could potentially alter a particular mutant's sensitivity to Stl. It has been well documented that many Rifampicin resistant RNAP mutants have altered termination and pausing activities (137-139). Specifically, Fischer and Yanofsky showed that the RpoB2 mutant has decreased termination and pausing activities, resulting in an overall

increase in transcription elongation rate. Conversely, the RpoB8 mutant has increased pausing and termination, resulting in an slower elongation rate (137). We decided to utilize the properties of these two mutants and test for changes in Stl sensitivity. As shown in Figure 17A, the EW1b *rpoB8* mutants were markedly hypersensitive to Stl treatment, while the EW1b *rpoB2* mutants were slightly resistant.

We made MG1655 *mrcB rpoB8* and *rpoB3595* double mutants to test if altering RNAP pausing frequency also affects sensitivity to Actinomycin D. RpoB3595 is a "fast" (less pausing, faster rate of elongation) polymerase with properties similar to RpoB2 (140). MG1655 *mrcB rpoB8* mutants were no more sensitive to ActD than wildtype, however the MG1655 *mrcB rpoB3595* mutants were very resistant to ActD treatment (Figure 17B). All *rpoB* mutants were confirmed with sequencing. The extreme level of ActD resistance caused by the *rpoB3595* mutation could possibly indicate that this mutant is capable of causing structural rearrangements that prevent ActD from inhibiting RNAP (see Discussion). These results confirm that the pausing frequency of the polymerase can affect sensitivity to elongation inhibitors.

Testing Rifampicin-resistant mutants for ActD sensitivity

Given that there has been no evidence for direct RNAP-ActD interactions, we were surprised to find that modulations in RNAP can lead to ActD resistance. We therefore screened additional rifampicin-resistant mutants for their effects on ActD

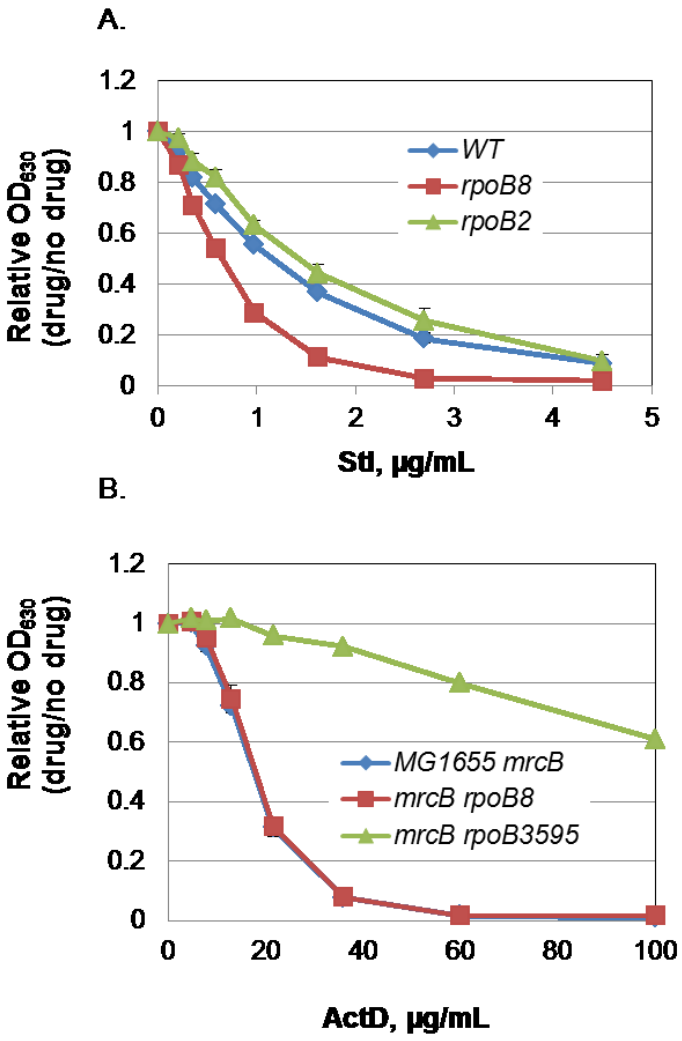


Figure 17: Drug sensitivity profiles for RNAP mutants with altered pausing frequencies

Hypersensitivity profiles for RNAP mutants with altered transcription properties were conducted as in Figure 2. In Panel A, RNAP mutants with point mutations in RpoB were tested for Stl sensitivity. EW1b *rpoB8* mutants are known to as "slow" polymerases, while EW1b *rpoB2* mutants are known to be "fast" polymerases. In Panel B, MG1655 *mrcB rpoB8* and MG1655 *mrcB rpoB3595* mutants were tested for ActD sensitivity. RpoB3595 is another "fast" polymerase mutant with similar properties as RpoB2. Error bars represent the standard deviation for three independent cultures.

sensitivity. MG1655 *mrcB* cells were plated on plates containing rifampicin, and 40 independent colonies were tested for ActD sensitivity. Each mutant was tested with our standard growth assay, and mutants with potential resistance were restruck and screened a second time. Those that passed both tests were sequenced and triplicates were tested in our quantitative assay (Table 4a and Figure 18 A-C). 20 of the 40 mutants were also sequenced regardless of sensitivity to get a rough sample of the mutations that do not lead to resistance (Table 4b). Out of the 40 mutants, we found 8 mutations in 6 positions that resulted in ActD resistance when tested in triplicate, and no mutants that were hypersensitive. Some of the mutants (including S574Y and Q148L) were isolated multiple times. The location of the mutations in the folded protein was determined by aligning the *E.coli* RNAP protein sequence with the *T. aquaticus* sequence and mapping the mutations to the crystal structure of *T. aquaticus* RNAP complexed with rifampicin. The mutations were located fairly evenly around the Rif binding pocket (Figure 19). One of the mutants, RifR22, clearly showed ActD resistance, yet grew inconsistently and did not show any mutation in the two typical locations of Rif-resistant mutations in *rpoB* or in all of *rpoC*.

Isolating ActD resistant mutants

We were also interested in identifying other RNAP mutations that lead to ActD resistance outside of the Rif binding pocket. Identification of additional mutations could shed light on the interactions between RNAP-ActD and further characterize the

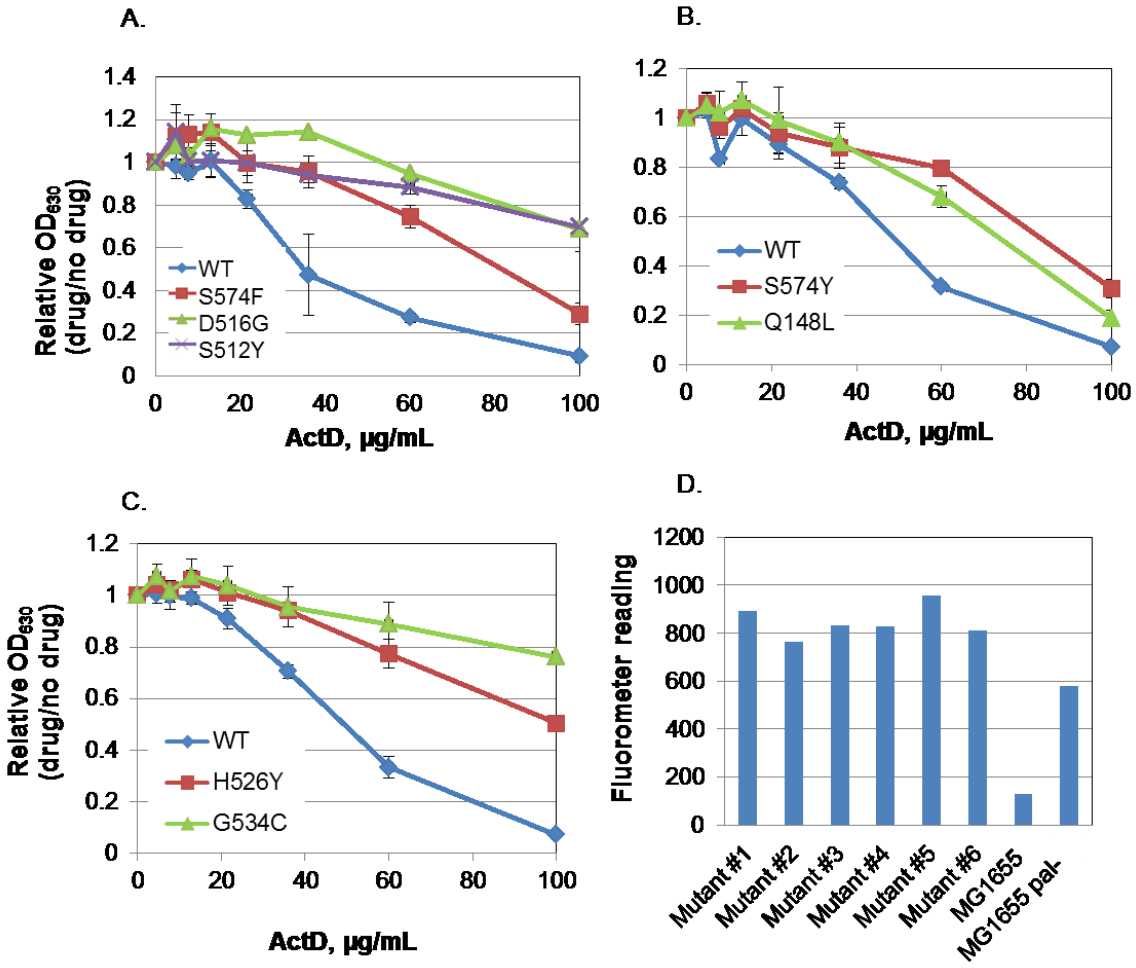


Figure 18: Identifying Actinomycin D-resistant mutants

In Panel A-C, the ActD sensitivity profiles for rifampicin-resistant mutants that are also ActD-resistant are shown. The mutants are listed by the mutation that leads to rifampicin and ActD resistance. Error bars are representative of the standard deviation for 3 independent cultures. In Panel D, six MG1655 *pal* ActD resistant mutants were tested for Sytox green incorporation. Wildtype MG1655 and MG1655 *pal* cells were included for comparison. Fluorometer reading is the average of two readings for the same sample.

Table 4a: RpoB mutations that lead to ActD resistance

Mutant #	Mutation	Carol Gross Nomenclature (140)
2	S574F	
5	I572F	RpoB7
18	D516G	
19	S512Y	
22	unknown	
21	S574Y	
26	S574Y[a]	
27	Q148L	
28	Q148L[a]	
30	H526Y	RpoB2
32	Q148L	
37	S574F[a]	
39	G534C	

[a] These mutants passed the first round of screening, but were not tested in triplicate due to redundancy

Table 4b: RpoB mutations that do not lead to ActD resistance

Mutant #	Mutation	Carol Gross Name
1	L511P	
23	Q513L	RpoB101
24	T563P	RpoB3370
33	D516N	RpoB113
34	I572T	
36	S574Y[b]	
38	H526Q	
40	S531F	RpoB 114

[b] This mutant passed the first round of screening but failed in the second

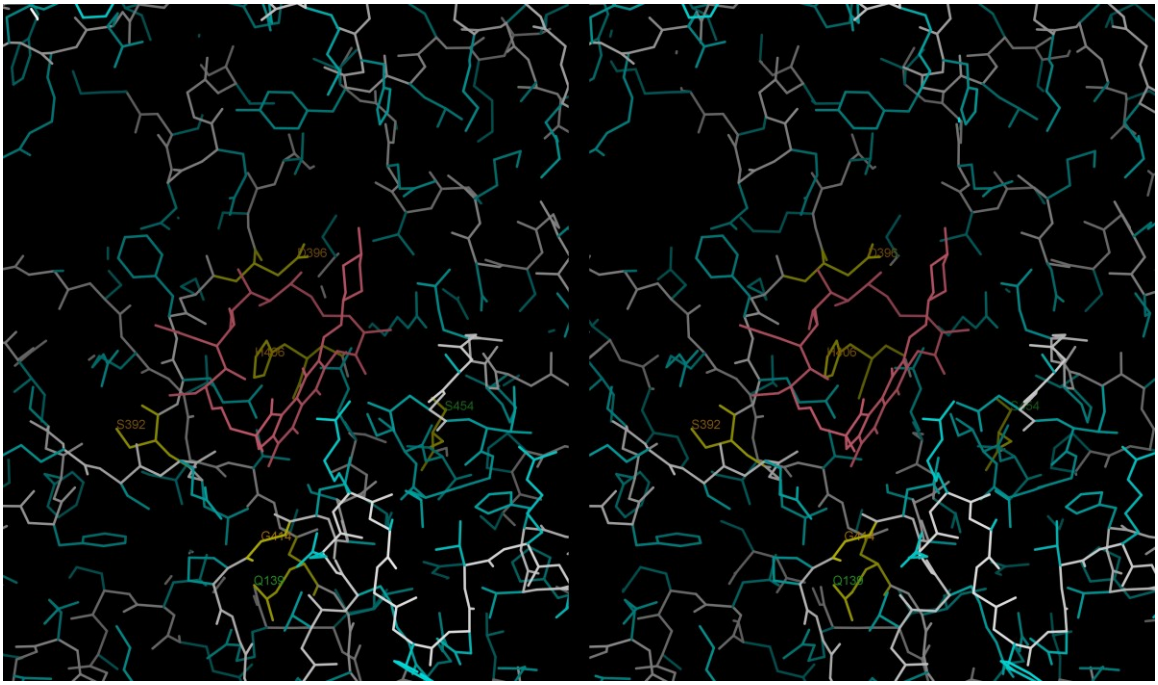


Figure 19: KiNG image of RNAP complexed with Rifampicin
This figure is a KiNG stereoimage of RNAP complexed with Rifampicin showing the position of the mutations that lead to ActD resistance. The RNAP backbone is shown in grey, the sidechains are shown in blue, the sidechains that are mutated in ActD resistant cells are shown in yellow, and rifampicin is shown in pink.

mechanism of ActD inhibition of elongation. We plated dilutions of MG1655 *pal* cells on both LB and ActD (25 μg/mL) to isolate ActD resistant colonies, and found a mutation frequency of 9.85×10^{-6}. However, when we sequenced both the *rpoB* and *rpoC* genes in several of the ActD resistant mutants, we found no mutations in either gene.

Given the high mutational frequency, we were concerned that ActD might itself be a mutagen. We took MG1655 *pal* cells that were either untreated, treated with 50 μg/mL ActD, or treated with 100 μg/mL ActD for 15 or 30 minutes and plated them on Rifampicin plates to compare mutational frequencies. The relative (drug/no drug) mutational frequencies were 0.5 and 0.6 for the 15 minute and 30 minute samples respectively for the 50 μg/mL ActD treated cells, and 1 and 0.4 for the 15 minute and 30 minute samples for the 100 μg/mL ActD treated cells. These results show ActD-induced mutagenesis is not the cause of the high frequency of ActD resistant cells.

Another model for the high number of ActD resistant cells is that there is a secondary mutation that is overcoming the *pal* mutation, which would restore the integrity of the cell membrane and prevent uptake of the drug. Sytox Green (Lifetechnologies) fluoresces when it intercalates into DNA, and cannot enter cells with intact membranes. To measure Sytox incorporation, we spun down overnights of six different ActD-resistant colonies, and MG1655 negative control, and a MG1655 *pal* control, resuspended in Tris-HCl, and treated with Sytox for 5 min in the dark. Fluorescence was measured with an excitation of 500 nm and an emission of 550 nm. As

shown in Figure 18D, fluorescence levels for each of the 6 mutants were much closer to MG1655 *pal* levels than MG1655, implying the membranes are still compromised. However, the possibility still exists that another mutation, such as one that increases the activity of a drug exporter, is modifying cell sensitivity to ActD in a way that does not explain the molecular mechanism of ActD inhibition on transcription.

3.3 Discussion

The antibiotics Streptolydigin and Actinomycin D are widely known for their ability to stall transcription elongation complexes. However, several key questions remain about the mechanism of their inhibition which are addressed in this chapter. Wildtype *E. coli* is resistant to both Stl and ActD, requiring modifications to be made to the cells for study in this host system. For our Stl studies we use the cell line EW1b, which has an insertion in the *tolC* gene and is therefore deficient in the efflux pump that normally pumps out Stl. For ActD studies, we found that deleting *pal* and *mrcB* in MG1655 cells makes the membrane more permeable to the drug.

The ability to remove Stl and resume elongation in *in vitro* studies implies that the binding is transient, but this has yet to be shown *in vivo* (59). One way to monitor the length of elongation stalling is the identification of replication-transcription collisions. Replication-transcription collisions can occur in healthy cells, as the replication machinery travels over 10 times faster than transcription. While co-

directional collisions do not impede replication, head-on collisions and those with stalled complexes lead to replication fork collapse, induction of the SOS damage response pathway, and cell death (38,39,70,81,141). Therefore, we tested cells treated with Stl for SOS induction using RecA western blots and found no changes in RecA levels with Stl treatment, supporting the model of transient Stl binding.

We also examined elongation inhibition of both drugs *in vitro.* In the Stl *in vitro* studies, Stl failed to inhibit RNAP at all levels tested, including those well above levels shown previously to inhibit RNAP *in vitro* when the substrate was coliphage T4 DNA. We believe that in our system transcription is too rapid to allow for Stl binding; presumably RNAP needs to be held in an "open" conformation for some period of time to allow for Stl binding. Stl binds near the active site in the center of the "crab claw" of RNAP. For this to occur, Stl most likely enters through the secondary channel, which is only accessible while RNAP is held in an open conformation. Rapidly elongating RNAP would only be in an "open" conformation for a very short period of time during the nucleotide addition cycle, while paused and backtracked complexes are held in the open conformation (Dr. Dorthy Erie, personal communication). A similar mechanism was shown for GreA activity, which is able to reactivate RNAP when it is kinetically trapped in an inactive [open] conformation (136).

In the ActD *in vitro* studies, 0.8 μM ActD was sufficient for nearly complete RNAP inhibition. However, we found that the titration window for ActD is very tight,

as ActD at 0.67 μM hardly impeded transcription. The ActD-DNA interaction, in particular with undistorted DNA, is quite slow to reverse (62); perhaps under our conditions once ActD is present at high enough levels the binding is essentially irreversible. Accordingly, we have been unable thus far to create conditions that allow for the resumption of elongation after initial ActD inhibition. Further studies will be necessary to be able to test Mfd release of RNAP stalled by either inhibitor *in vitro*.

Given our proposal that Stl binding requires RNAP to be held in an open conformation, such as at a pause site, we hypothesized that altering the pausing frequency of RNAP could affect Stl and ActD sensitivity. We looked at RNAP mutants known to have altered elongation properties and showed that the increased pausing frequency can cause hypersensitivity to Stl treatment, while decreased pausing frequency can cause resistant to both Stl and ActD. This further supports our suspicions that Stl binds more efficiently to paused RNAP.

We also screened 40 Rif-resistant colonies to expand on the finding of a RpoB mutant that is resistant to ActD. We found 8 mutations in 6 locations that lead to ActD resistance, further confirming the role RNAP structure can play in causing ActD resistance. The extreme resistance of several of the *rpoB* mutants to ActD is not likely due to a simple decrease in pausing frequency; rather, it is more consistent with ActD being nearly completely unable to inhibit these RNAPs. We hypothesize that these mutations cause structural rearrangements in the RNAP-DNA-ActD complex. These

structural rearrangements could potentially be alterations to the DNA structure in the transcription bubble, as it has been proposed that ActD binds to distorted DNA at the ends of the transcription bubble (64). ActD binding and release rates are increased with DNA distortion (62), and possibly the DNA in the transcription bubble is distorted in such a way that ActD does not stably bind to DNA.

In an alternative model, if RNAP and ActD interact directly, the *rpoB* mutations could cause structural rearrangements to the ActD-RNAP interface. This model resembles the inhibition mechanism of topoisomerase inhibitors, which bind to both specific DNA targets and the topoisomerase. *In vitro* mutational studies showed that inhibitors are highly specific for the DNA sequences immediately flanking the cleaved DNA, and photocrosslinking experiments showed 3-azido-AMSA crosslinks to the base-pairs flanking the cleaved bonds (142,143). Additionally, mutations overlapping the active site in the T4 topoisomerase genes *gp39* and *gp52* alter sensitivity to both quinolones and antitumor agents, implying that these inhibitors bind to a conserved site in topoisomerase (144). ActD is well-known as a DNA intercalator, yet it could also directly interact with RNAP near the Rif binding site.

We set out to identify other RNAP mutants that are resistant to ActD outside the Rif binding site in an attempt to separate these two models. We easily isolated ActD-resistant mutants, and found a surprisingly high mutation frequency leading to ActD resistance. While we showed that the membrane of the mutant cells were still

compromised, the possibility still exists that these mutations are causing up-regulation of a transporter system that is pumping ActD back out of the cells. This could be similar to the case of Stl; *E. coli* cells treated with EDTA presumably can uptake Stl due to a more permeable membrane, but the drug is pumped backed out via the TolC system. *tolC* mutants are resistant to ActD, but other export systems exist that could be producing the resistance phenotype. A greater understanding of the high mutation rate is necessary before attempting to elucidate the mechanism of ActD inhibition by mutational screening.

3.4 Materials and Methods

Materials

Luria–Bertani broth (LB) contained Bacto tryptone (10 g l-1), yeast extract (5 g l-1), and sodium chloride (10 g l-1) and was used for all bacterial growth (with appropriate antibiotics for plasmid selection and the indicated additions); Nitrocellulose membrane (Protran® BA 85) was from Whatman; Streptolydigin was the generous gift of Konstantin Severinov (Waksman Institute of Microbiology) and José Salas (Universidad de Oviedo; Actinomycin D was ordered from Sigma; RecA antibodies were purchased from Abcam. RNA polymerase was kindly provided by Dorthy Erie (University of Chapel Hill).

***E. coli* strains:**

EW1b [F-, lacY1 or lacZ4, tsx-64, *glnV44*(AS), gal-6, *LAM-*, *hisG1*(Fs), *DtolC5*, *argG6, rpsL* (allele 8, 104 or 17), *malT1*(LamR)] was obtained from the *E. coli* Genetic Stock Center (Yale University). MG1655 [F- lambda- *ilvG- rfb-50 rph-1*] was obtained from the E. coli Genetic Stock Center (Yale University). MG1655 and EW1B derivatives were constructed by phage P1-mediated transduction from the Keio collection (114). The *rpoB8*, *rpoB2*, and *rpoB3595* mutants were P1 transduced from the Carol Gross collection (140). Mutations were confirmed with PCR and sequencing.

Growth kinetics for Stl and ActD sensitivity

Overnight cultures in LB were diluted to roughly $4 * 10^6$ cell/ml in LB, and 75 ml was delivered to each well in a microtiter plate. 75 ml of Stl or ActD at twice the indicated concentration (or drug-free control) was also added to each well, for a total volume of 150 ml per well. The plate was incubated at 37°C for 12 h with constant shaking in a BioTek ELx808 Microplate Reader. The optical density (at 630 nm) of each well was read every 15 min.

Western Blots

Quantitative RecA Western blots were analyzed in EW1b and EW1b *mfd* cells. Cells were pregrown overnight in LB media at 37 °C then diluted to an OD_{560} of 0.1. The

cells were grown to OD_{560} 0.5 and treated with 5 μg/mL Stl or LB for 1 hour, when 8 x 10^8cells were collected for each sample. Cell pellets were stored at -20 °C.

Pellets were thawed at room temp and resuspended in 100 μl of 20 mM potassium phosphate pH 7.4, 1 mM EDTA, and 0.5 mM DTT. 20 μL was diluted 1/5 om 0.5% SDS and boiled for 10 minutes. Protein concentration in the cell lysates were measured using BCA™ Protein Assay Kit (Thermo Scientific) against BSA standards. The amount of protein loaded per gel was equal to 24 μL of the most dilute sample.

Samples were loaded onto a 7.5% polyacrylamide (Tris-HCl) gel and run for approximately 2 h in 25 mM Tris-Glycine buffer containing SDS (0.1%). Known amounts of RecA protein were also loaded to create a standard curve, allowing for quantitation of RecA levels for each sample

The portion of the gel containing proteins larger than about 75 kDa was cut off and stained with Coomassie blue dye to serve as a loading control. The remaining portion of the gel was transferred to a nitrocellulose membrane for 60 min at 12 V using a Genie Blotter transfer device (Idea Scientific Co.). The blot was blocked for 1 h in 20% non-fat milk powder solution (Biorad) in Tris-buffered saline (TBS). The membrane was incubated overnight at 4°C with polyclonal RecA primary antibody and Tween (0.1%), and then washed three times with TBS buffer at room temperature (10 min each). The membrane was incubated with secondary antibody (IRDye 800CW-conjugated goat anti-rabbit IgG (LI-COR®)) for 1 hour, and the washes were repeated. After air-drying, the

membrane was scanned on an Odyssey Imaging System (LI-COR Biosciences), and quantified using the Odyssey software.

***In vitro* transcription reactions**

For *in vitro* reactions with Stl, open complexes were formed by adding 20 nM DNA template (254-bp PCR fragment from pDE13 containing the λ_{PR} promoter and t_{R2} terminator) and 20 nM RNAP in 1x transcription buffer (80 mM Tris-HCl pH 8.0, 200 mM KCl, 16 mM $MgCl_2$, 200 uM EDTA pH 8.0, 1 mM DTT, 0.05 mg/mL BSA and 16% Glycerol). Open complexes were incubated at 37 °C for 10 minutes. Elongation to +24 was carried out by adding NTPs (200 μM ATP, 0.3 μM ^{32}P– GTP, and 15 μM UTP) and heparin in 1X transcription buffer and incubating at room temp for 5 min. Stalled complexes were split into 25 μL each and reactions were treated with 200 μM of UTP, GTP, and CTP, 4 mM ATP, the indicated amount of Stl, and either 400 nM Mfd or the equivalent volume of Mfd storage buffer and incubated at 37 °C for 5 min.

For the reactions containing ActD, +24 complexes were set up in the same manner as the Stl reactions and split into 25 μL reactions. Complexes were treated with the indicated amount of ActD or 1x transcription buffer and incubated at 37 °C for 10 min. 200 μM of each NTP was added and complexes were incubated for 1 min at room temp.

All reactions were quenched with 50% formaldehyde and 10x native loading dye and placed on ice until being loaded in the gel. Samples were loaded into 8 M urea, 20%

polyacrylamide native gels and run at 100 V until the loading dye had run about 9/10 of the gel length. The bottom portion of the gel (containing the majority of the free nucleotides) was excised and the gel was exposed to a phosphoimager screen overnight.

Actinomycin D mutagenesis test

MG1655 *pal* cells were pregrown overnight in LB and diluted to approximately OD_{560} 0.1. Cells were grown in LB to OD_{560} 0.5, when they were treated with either LB, 50 μg/mL ActD, or 100 g/mL ActD. 2 mL of an untreated control sample was taken before drug treatment. 4 x 10^8 cells were collected at 15 and 30 minutes, spun down, and plated onto plates containing rifampicin (100 μg/mL).

Sytox Green test for membrane integrity

MG1655 *pal* cells were pregrown overnight in LB and plated on LB plates containing Actinomycin D (25 μg/mL). Six resulting ActD-resistant colonies were pregrown overnight. 1 mL of overnight culture was spun down and resuspended in 600 μL of Tris-HCl. Sytox Green (Lifetechnologies; 3 mM) was added to the cells for 5 min in the dark. Fluorescence was measured with an excitation of 500 nm and an emission of 550 nm.

4. Fate of transcription elongation complexes stalled by exogenous elongation inhibitors

Some of this work was published in Molecular Microbiology (84). The work for the Actinomycin D portion of this chapter was assisted by Ramsey Al-Khalil, an undergraduate researcher in our lab.

4.1 Introduction

With the increased understanding of Stl and ActD inhibition of transcription from Chapter 3, we are able to address the cellular response to RNAP stalled by Stl and ActD (46,59). This allows us to isolate the effects of stalled transcription, eliminating many of the secondary effects that complicate the studying of blocked transcription using DPCs. Therefore, the primary goal of this Chapter is to understand how cells clear elongation complexes blocked by Stl and ActD.

While it has been shown that Stl-RNAP complexes are stable *in vitro* and can be reactivated with the removal of Stl (59,75), the fate of elongation complexes stalled by Stl *in vivo* is currently unknown. In contrast to the *in vitro* studies, von Meyenburg and colleagues showed that elongation complexes are inactivated upon Stl treatment *in vivo*. To monitor the effect of Stl on elongation complexes, they treated cells with Rif and Stl and measured both RNA and protein synthesis. Treatment with Stl (100 μg/mL) resulted in a halt in RNA synthesis, with removal of the drug resulting in a resumption of elongation after a delay. Surprisingly, when these conditions were repeated with the addition of Rifampicin after the removal of Stl, there was no new RNA and protein

synthesis; instead RNA and protein levels decreased slightly. This implies that the Stl-stalled complexes are unable to continue elongation even when the drug is removed, and in the absence of Rif, elongation only continues due a new round of initiation. It was this result that led them to suggest that Stl treatment renders RNAP susceptible to interactions with an unknown termination factor, leading to release of the complex (58). To date this termination factor remains unknown, and no additional evidence has been presented. Therefore, we set out to determine the mechanism of RNAP release during Stl treatment.

We show here that *mfd* knockout cells are slightly resistant to Stl treatment, and Mfd-overexpressing cells are hypersensitive, implicating the transcription-coupled repair pathway in recognizing Stl-stalled RNAP. It is of interest to see if *mfd* mutants are also resistant to ActD for two reasons: 1) To see if Mfd activity is general to multiple transcription inhibitors and 2) To see if the model presented in *E. coli* could be relevant for eukaryotes. The second point is of particular interest given that certain platinum resistant tumors have been shown to have elevated levels of CSB, the eukaryotic homolog to Mfd, (145) and could potentially be hypersensitive to ActD.

In this Chapter, we propose a model in which Mfd-mediated release of RNAP stalled by Stl or ActD is in balance with natural elongation that can occur once the drug binding is reversed. Overexpression of Mfd results in a tipping of the balance towards

premature termination, leading to cell death. These results are consistent for both Stl and ActD, implying that they could be relevant for eukaryotic systems as well.

4.2. Results

The transcription coupled repair pathway recognizes Stl-stalled RNAP

von Meyenburg *et al.* showed that Stl-stalled RNAP is released *in vivo* but did not identify the protein responsible (58). We previously hypothesized that the transcription-coupled repair factor Mfd was the release factor, and a previous lab member showed *mfd* mutants are partially resistant to Stl (84). These data were confirmed using the drug sensitivity profile tests used in previous chapters (Figure 20A). We were also interested in testing cells overexpressing Mfd, and found that increasing Mfd activity by leads to hypersensitivity to Stl (Figures 20B) (84). These results were surprising given Mfd's currently understood role as a beneficial repair protein, as removal of Mfd typically hinders cell survival under certain stressful conditions.

We were then interested in testing *mfd* mutants with the eukaryotic transcription inhibitor Actinomycin D to see if the results remain consistent for multiple inhibitors.

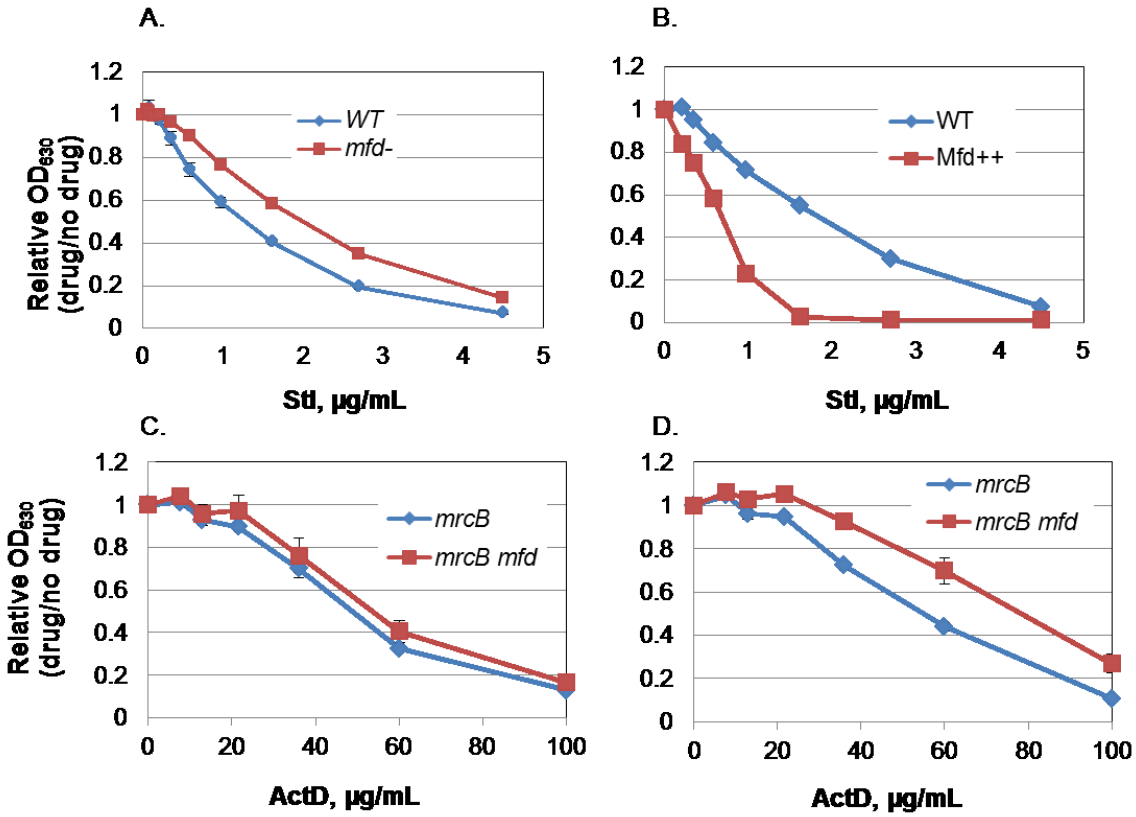


Figure 20: Effect of Mfd on cell growth during Stl and ActD treatment .
In Panel A, three independent cultures of EW1b (WT) or EW1b *mfd* were tested for Stl sensitivity using a microtiter plate sensitivity assay as in Figure 2. In Panel B, EW1b cells carrying pCA24N-Mfd, which expresses Mfd on an IPTG inducible promoter, were tested for Stl sensitivity as in Panel A. In Panels C and D, MG1655 (WT) and mutant derivatives were tested for ActD sensitivity as in Panel A. Panels A-C have an endpoint of 50% of the maximum growth rate, while Panel D is the comparative titration curve for the same cells as in Panel C using 10% of the maximum growth rate as the endpoint. Error bars for all panels are representative of the standard deviation for three independent cultures.

MG1655 *mrcB mfd* double mutants were constructed using the Keio collection and tested for ActD sensitivity using the same hypersensitivity profile test. When using an end timepoint of 50% of the maximum growth rate the MG1655 mrcB mfd cells were only slightly resistant to ActD treatment (Figure 20C). However, by the end of exponential growth the MG1655 *mrcB mfd* cells were clearly resistant compared to MG1655 *mrcB* cells (Figure 20D), supporting a more general role for Mfd in recognizing RNAP stalled by elongation inhibitors.

Candidate mutant approach to look for other proteins that recognize Stl-stalled RNAP

There are two potential models for the resistance of *mfd* cells: either release of stalled RNAP complex is needed for growth and Mfd is interfering with the release pathway, or Mfd is releasing RNAP and this activity is detrimental to cell growth. In the latter case, the transient pausing demonstrated in Chapter 3 indicates that Mfd would be releasing polymerases that would otherwise be able to complete transcription, resulting in detrimental premature termination. To separate these two possibilities, we first need to identify any other proteins that could potentially respond to Stl-stalled elongation complexes. We first looked at transcription terminator Rho using the same double-drug titration as with Aza-C (see Chapter 2) and found that reducing transcription terminator Rho activity using bicylomycin does not result in increased sensitivity to Stl (data not shown) (84). We then tested several factors known to affect transcription elongation and

termination: GreA, GreB, HepA, DinG, UvrD, Rep, and DksA(see Introduction in Chapter 2). YejH was also added to this list, as it is a putative helicase that, when overexpressed, can complement UV damage repair in *mfd* deficient cells (146). We also tested double mutants to look for functional redundancy. *uvrD rep* double mutants are lethal, presumably due to blockages to replication that are normally cleared by these proteins, but the addition of a *recF* mutation can restore viability ((71) and references therein).

Using the same drug sensitivity profile test we found no mutants that were strongly hypersensitive to Stl treatment (Table 5). Some single and double mutant strains, however, were slightly resistant to Stl, including *dksA, rep, dksA mfd, mfd yejH, dinG rep,* and *rep mfd* mutants. *dksA* cells were resistant only at low levels of Stl, while *dksA mfd* cells appear to have an additive resistance. *rep* cells were only very slightly resistant, and the *dinG rep* and *mfd rep* cells were resistant at the same level as the *rep* single mutant (Figure 21 and data not shown). The *mfd yejH* strains were also only as resistant as the *mfd* single mutant. Therefore it appears that the proteins that consistently recognize Stl-stalled RNAP are Rep and Mfd, with DksA influencing cell growth at low levels of Stl treatment.

We also attempted a screen for Stl hypersensitive and resistant mutants to obtain a comprehensive list of mutants that alter cell sensitivity to Stl. However, this screen

Table 5: Stl hypersensitivity profile results for transcription elongation and termination factors

Strain	Growth compared to WT in the presence of Stl
greA	No change
Rho (Bcm)	No change
dinG	No change
uvrD	No change
yejH	No change
dinG, uvrD	No change
uvrD/recF/rep	No change
rep	Slightly Resistant
dinG/rep	Slightly Resistant
dksA	Resistant (low levels of Stl)
mfd	Resistant
dksA/mfd	Resistant
yejH/mfd	Resistant
rep/mfd	Resistant

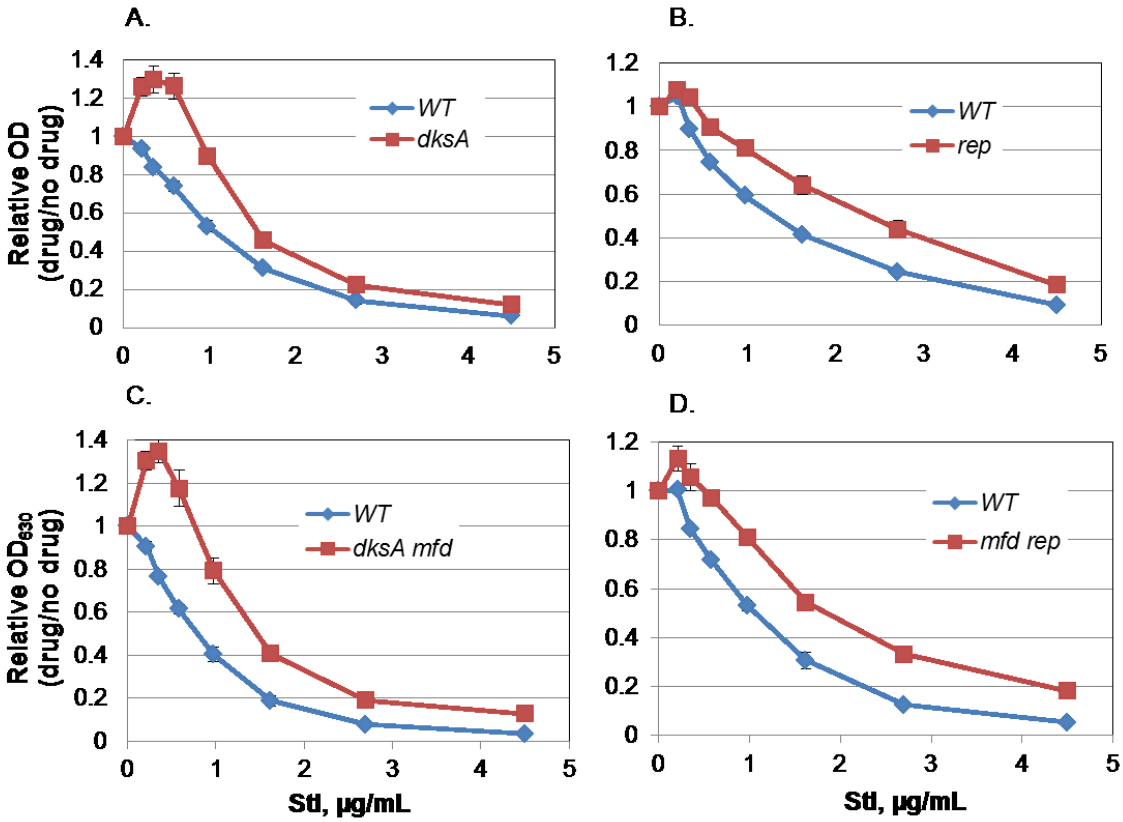


Figure 21: Stl hypersensitivity profiles of mutants resistant to Stl.
Three independent cultures of EW1b (WT) or mutant derivatives were tested for Stl sensitivity using a microtiter plate sensitivity assay as in Figure 2, with error bars indicating the standard deviation of each. The endpoint OD was taken at 50% of the maximum growth rate.

proved to be technically challenging due to the resistance of normal bacteria to Stl and was later abandoned in favor of other directions.

Mfd activity creates substrates for the tmRNA system

In Chapter 2 we showed that the tmRNA system is vital for cell survival during DPC formation. This is due to the necessity of recycling the stalled ribosomes, since the coupling of transcription and translation results in both process being stalled at DPCs. Therefore, treatment with Stl is also expected to block translation, and cells could require the presence of the tmRNA system during Stl treatment. To verify the requirement of the tmRNA system for cell survival we treated *smpB* knockouts with Stl. SmpB is an essential cofactor to the tmRNA system that assists with tmRNA binding to the ribosome. We found that *smpB* knockouts were hypersensitive as expected (Figure 22A) (84).

If Mfd is responsible for releasing Stl-stalled polymerases, the release of the nascent transcript would allow translation to continue to the end of the transcript and open up the A-site of the ribosome for tmRNA binding. It then follows that Mfd is responsible for creating the substrates for the tmRNA system, and removing Mfd would therefore remove the necessity of the tmRNA system as there would be no increase in substrates during Stl treatment. We therefore constructed *smpB mfd* double mutants to test their sensitivity to Stl. As shown in Figure 22C, *smpB mfd* mutants lose the

hypersensitivity of a *smpB* single mutant. This implies that Mfd is capable of creating substrates for the tmRNA system. However, the double mutant was also less resistant to Stl than a *mfd* single mutant, suggesting the tmRNA is capable of contributing to cell survival even in the absence of Mfd. As a comparison we constructed a *smpB rep* double mutant, since *rep* cells were also slightly resistant to Stl treatment. In contrast to the *smpB mfd* cells, the *smpB rep* double mutants were just as hypersensitive to Stl as a *smpB* single mutant, implying that *rep* cells are resistant due to an activity unrelated to RNAP release (Figure 22D).

To confirm that Mfd activity directly results in increased tmRNA tagging, we treated cells overexpressing Mfd with Stl and conducted western blots looking for tmRNA tagging. To overexpress Mfd we used EW1b cells containing pCA24N-Mfd, which overexpresses MFD from an IPTG inducible promoter. Cells overexpressing Mfd had greatly increased levels of tmRNA tagging compared to cells containing a vector plasmid during Stl treatment (See Figure 23), confirming Mfd mediated termination is creating substrates for the tmRNA system. We also were surprised to find that overexpression of Mfd, even in the absence of Stl treatment, led to tmRNA tagging. This indicates that Mfd has termination activity in wildtype cells, which is addressed in Chapter 5.

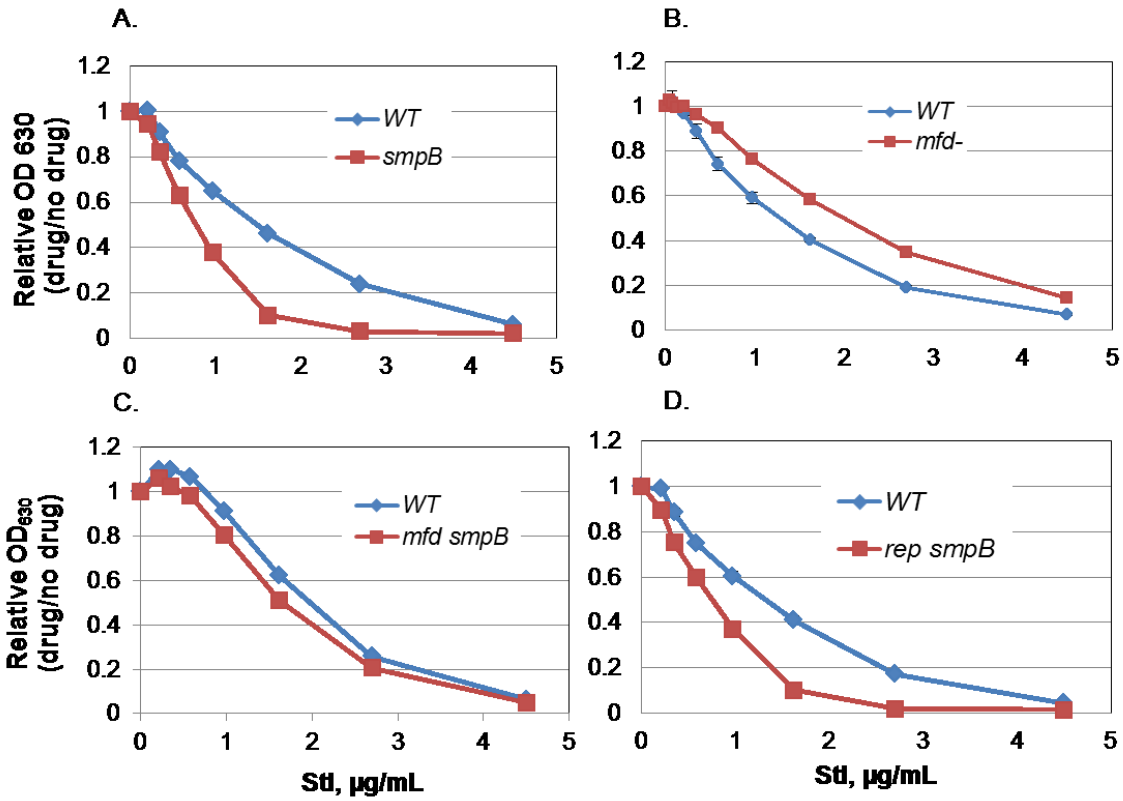


Figure 22: Stl hypersensitivity profiles for *smpB* mutants
EW1b (WT) or mutant derivatives were tested for Stl sensitivity as in Figure 2. Three independent cultures were measured for each strain. The error bars indicate the standard deviation for these cultures. *mfd* single mutant data is from Figure 20, shown again here for easy comparison. The loss of hypersensitivity in the *mfd smpB* double mutant compared to the *smp* single mutant indicates a loss of substrates for the tmRNA system. However, the loss of resistance in the *mfd smpB* double mutant compared to the *mfd* single mutant shows that tmRNA is capable of contributing to cell survival even in the absence of Mfd.

tmRNA	-	+	+	+	+	+	+	+	+	+	+	+	+
IPTG	-	-	-	-	+	+	+	-	-	-	+	+	+
Stl (5 ug/mL)	-	-	-	-	-	-	-	+	+	+	+	+	+

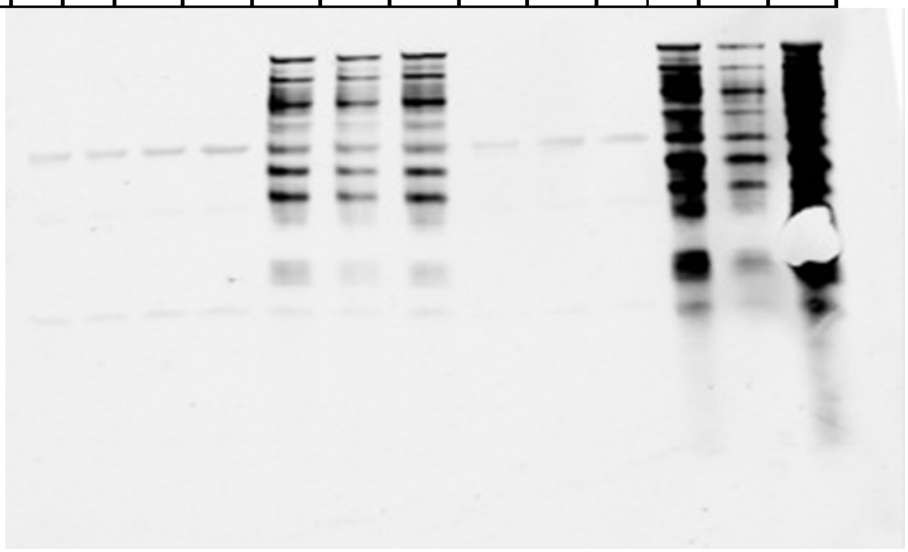


Figure 23: tmRNA tagging in Mfd-overexpressing cells
EW1b cells containing pCA24N-Mfd, which encodes Mfd on an IPTG inducible promoter, were grown to OD_{560} 0.5 with or without IPTG. They were then treated with Stl at 5 μg/mL (or left untreaded) for one hour. Western blots were conducted with antibodies against the degradation tag encoded by tmRNA. EW1B cells are protease proficient, so any tagging shown here has overwhelmed the protease system.

The hypersensitivity of *rpoB8* mutants to Stl is partially due to Mfd-induced termination

We found in Chapter 3 that *rpoB8* mutants, which have increased pausing frequency, are hypersensitive to Stl. The hypersensitivity of the *rpoB8* mutant could potentially be due to the decreased elongation rate of the polymerase; the longer the RNAP is on the template, the more time Stl has to bind and slow the polymerase down further. It could also be due to increased termination by Mfd. To test these different models, we constructed a *mfd rpoB8* mutant and compared it to a wildtype cells and *mfd* and *rpoB8* single mutants. As shown in Figure 24, the *mfd rpoB8* double mutant is slightly more sensitive than the *mfd* single mutant, implying that the slow speed of the RpoB8 mutant is enough to cause some sensitivity to Stl. However, the *mfd rpoB8* double mutant is significantly more resistant to Stl than the *rpoB8* single mutant, implying that Mfd termination is also responsible for some of the hypersensitivity of the *rpoB8* mutant. This confirms that both the pausing frequency of the polymerase and Mfd termination contribute to overall transcription rates during Stl treatment.

***In vitro* reactions with elongation inhibitors and Mfd**

In order to confirm that the *in vivo* data is truly the result of Mfd release of inhibitor-stalled RNAP, efforts were made to recreate this system *in vitro.* The end goal was to compare the amount of full length product in the presence and absence of Mfd. The first control test was to verify that Mfd releases nucleotide-starved RNAP (similar to

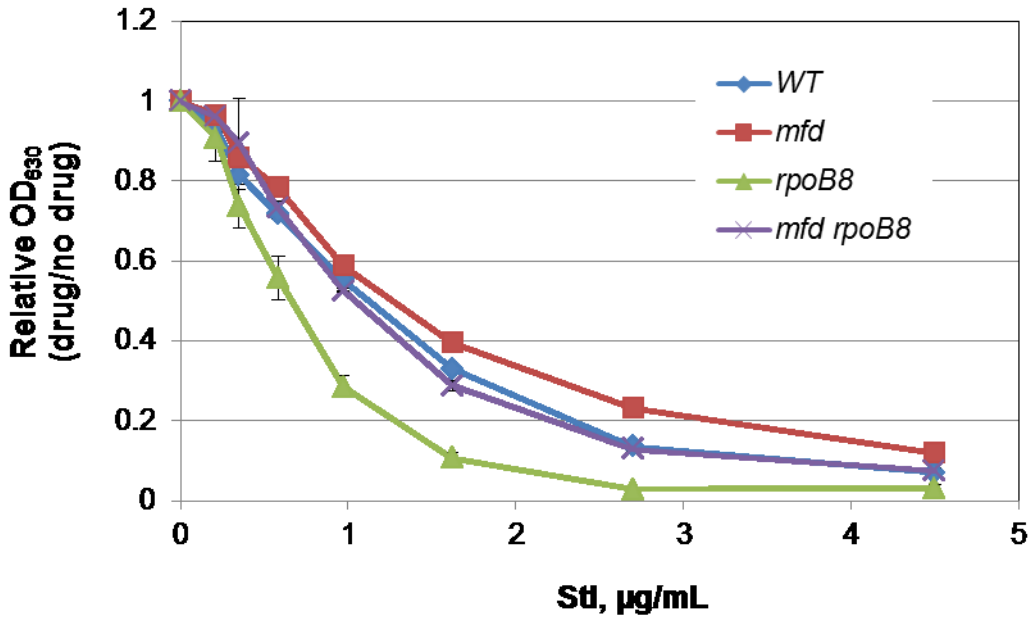


Figure 24: Drug sensitivity profiles monitoring the effect of Mfd activity on RpoB mutants

Hypersensitivity profiles for RNAP mutants with altered transcription properties were conducted as in Figure 2. The effect of Mfd on the slow RpoB8 mutant was tested using a *mfd* RpoB8 double mutant. Error bars are representative of the standard deviation for three independent cultures.

(147)). As shown in Figure 25A, nucleotide-starved elongation complexes are able to form full length transcripts when all four nucleotides are added, but not if they are preincubated with Mfd and high levels of ATP. This implies Mfd releases stalled RNAP in an ATP-dependent manner (requiring more than 200 μM ATP) during the preincubation step, and therefore the full length complexes do not form when all four NTPs are added.

We next needed to verify that addition of all four nucleotides at the same time as Mfd would allow for full length transcript production. Once RNAP is elongating Mfd cannot release it, as it must have an opposing force to dislodge the stable RNAP-DNA interactions. However, we were concerned that Mfd might release the stalled RNAP before elongation could resume. As shown in Figure 25B, resumption of elongation out-competes Mfd activity. We also noticed that there is significantly less +24 transcript in the Mfd treated samples. Perhaps this is indicating that there are small amounts of RNase present in the samples, and Mfd release of the +24 transcripts results in them being more vulnerable to degradation. In the absence of Mfd, the RNA would still be attached to the ternary complex and possibly protected from this degradation.

Even though it appeared Stl is not inhibiting elongation, if Stl is halting RNAP temporarily, Mfd might be able to terminate those complexes. Therefore we tested a titration of Stl with Mfd present. As shown in Figure 16, addition of Mfd did not result in an accumulation of intermediate length transcripts. Further attempts to stall RNAP,

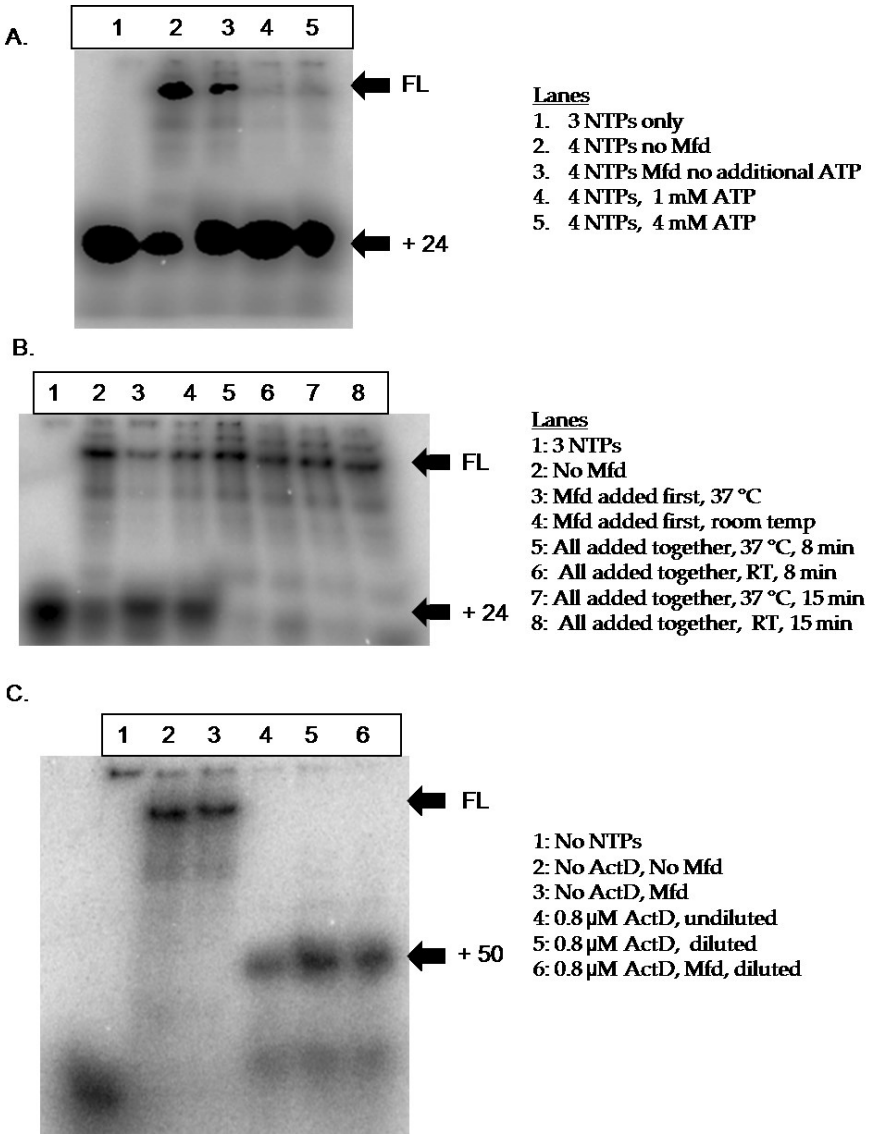


Figure 25: *In vitro* assay for Mfd-mediated release of Stl-stalled RNAP
Panel A is an *in vitro* assay monitoring Mfd release of nucleotide-starved RNAP. Transcription was carried out using radioactive GTP, which allows for visualization of resulting RNA products. +24 = product resulting from a nucleotide stalled RNAP at +24 of the transcript. FL= full length RNA product. The addition of Mfd and high ATP levels results in a loss of full length products. Panel B is a control for the competition between productive elongation and Mfd activity. The presence of full length products in lanes 5-8 indicate that productive elongation occurs before Mfd can release nucleotide starved RNAP. In Panel C the ability of Mfd to release ActD-stalled complexes was tested. Reactions containing stalled complexes were diluted to create sub-inhibitory levels of ActD, yet full length products were unable to form even in the absence of Mfd.

such as by introducing a strong pause site and preincubating with Stl, would have to be made for completion of these studies.

In Chapter 3 we showed that ActD treatment *in vitro* apparently completely inhibits elongation at 0.8 μM, while is virtually ineffective at 0.67 μM. To be able to measure Mfd activity, we need to have an intermediate level of ActD treatment that causes some stalling (creating substrates for Mfd), yet allows for full length production if Mfd does not release RNAP. Therefore we tried treating with inhibitory levels of ActD to cause stalling, then diluting to sub-inhibitory levels to allow for release. As shown in Figure 25C, dilution was not sufficient to cause release of ActD (see Lane 5, where the diluted should have been able to produce full-length transcripts if ActD was released). Further experiments are thus necessary to monitor Mfd activity (see Discussion).

4.3 Discussion

The data presented here provide evidence for a model in which the transcription-coupled repair pathway leads to premature termination of transcription stalled by elongation inhibitors. We have shown that knocking out Mfd leads to resistance to Stl or ActD treatment, while overexpressing Mfd leads to hypersensitivity to both inhibitors. This is in direct contrast to Mfd's well-characterized role as a beneficial repair protein, where Mfd activity is necessary for cell survival during DNA damage (such as with UV treatment).

The hypersensitivity of Mfd-overexpressing cells to Stl and ActD opens up intriguing possibilities for use of ActD in eukaryotic combination therapies. Eukaryotic transcription coupled repair is similar to that of bacteria, yet the specific mechanisms and proteins involved are more complicated. In yeast, the transcription coupled repair pathway is primarily mediated by Rad26, the yeast homologue of Mfd. It has been proposed that the helicase-like activity of Rad26 can help alter the structure of RNAPII, allowing nucleotide excision repair to fix the damage and RNAPII to resume productive elongation. When RNAPII is left permanently trapped Rad26 forms a complex with Def1 and induces degradation of RNAPII by ubiquitin-mediated degradation. Similar to the dual activities of Mfd, Rad26 alone appears to protect RNAPII from Def1-dependent degradation, preventing detrimental termination of elongation (148). It has also been shown in human cells that RNAPII degradation occurs after UV treatment, and that this degradation is dependent on the human transcription coupled repair factors CSA and CSB (149), although the exact fate of RNAPII during transcription-coupled repair in eukaryotes is still under debate. Furthermore, certain platinum-resistant tumors have been shown to have elevated levels of CSB (145). It would be interesting to see if ActD treatment stimulates RNAPII degradation in these platinum-resistant tumors in a similar manner as Stl treatment stimulates release of RNAP in Mfd-overexpressing *E. coli* cells.

To verify that Mfd is the main response factor to these inhibitors, and not a competitor to a beneficial repair pathway, we used a candidate mutant approach and

tested the Stl sensitivity to known transcription elongation and termination factors. No mutants tested were hypersensitive, and only *dksA* and *rep* cells showed slight resistance to Stl. DksA is an elongation factor that does not appear to have any RNAP releasing activity; instead DksA has shown to promote RNAP elongation (81). Rep is known to remove proteins bound to DNA ahead of the replication forks, and specifically has been shown to remove RNAP during transcription/translation collisions (70,71). We propose Rep is removing a small fraction of stalled RNAP ahead of replication forks. This activity does not apparently create substrates for the tmRNA system, suggesting that the stalled ribosomes have already been released (see below). It is also interesting to note that while *rep* single mutants and *rep dinG* mutants are resistant, the *uvrD rep recF* triple mutant is sensitive to Stl at wildtype levels.

In Chapter 2 we identified the involvement of the tmRNA system in cell survival with aza-C treatment, but were unable to identify the factor(s) that allow tmRNA to bind to the A-site of the stalled ribosome. In this chapter we showed tmRNA also assists in cell survival during Stl treatment, and that Mfd is the release factor responsible for the majority of tmRNA activity by releasing RNAP and the nascent transcript. Presumably, the stalled ribosome continues translating until the end of the nascent transcript, where the A-site would be left empty and tmRNA can act. The fact that *smpB mfd* mutants were more sensitive to Stl than *mfd* single mutants suggests that tmRNA is still capable

of releasing some stalled ribosomes in an Mfd-independent manner. This activity could be precluding Rep's removal of RNAP ahead of the stalled ribosome.

We found in Chapter 3 that *rpoB8* mutants, which have increased pausing frequency, are hypersensitive to Stl. This is presumably due to Stl binding more frequently to the paused polymerases. However, we also show here that part of the hypersensitivity of this mutant to Stl is due to the increased time for Mfd to release RNAP, providing further evidence for the model where Mfd termination is in competition for normal elongation.

Taken together with the previous data, we proposed the following model for Stl-induced stalling of elongation. In the absence of Mfd activity, Stl treatment stalls RNAP only transiently, and normal elongation can continue. Because the A-site is never unoccupied, tmRNA is unable to access the stalled ribosome, making the tmRNA system dispensable in *mfd* cells. However, when Mfd is present some of the Stl-stalled RNAP is released before Stl binding reverses. tmRNA then recycles the stalled ribosome and tags the peptide for degradation. When Mfd is overexpressed, Mfd out-competes productive elongation and the cells become hypersensitive.

4.4 Materials and Methods

Materials

Luria–Bertani broth (LB) contained Bacto tryptone (10 g l-1), yeast extract (5 g l-1), and sodium chloride (10 g l-1) and was used for all bacterial growth (with appropriate antibiotics for plasmid selection and the indicated additions); Nitrocellulose membrane (Protran® BA 85) was from Whatman; Streptolydigin was the generous gift of Konstantin Severinov (Waksman Institute of Microbiology) and José Salas **(**Universidad de Oviedo); Actinomycin D was purchased from Sigma; RNA polymerase was kindly provided by Dorthy Erie (University of Chapel Hill); ssrA antibodies were a generous gift from Tania Baker (MIT).

***E. coli* strains**

EW1b [F-, lacY1 or lacZ4, tsx-64, *glnV44*(AS), *gal-6*, LAM-, *hisG1*(Fs), *DtolC5*, *argG6*, *rpsL* (allele 8, 104 or 17), *malT1*(LamR)] was obtained from the *E. coli* Genetic Stock Center (Yale University). MG1655 [F- lambda- *ilvG- rfb-50 rph-1*] was obtained from the *E. coli* Genetic Stock Center (Yale University). MG1655 and EW1B derivatives were constructed by phage P1-mediated transduction. Knockouts of *mrcB*, *greA*, *greB*, *hepA*, *dksA*, *dinG*, *uvrD*, *rep*, *yejH*, and *smpB* were P1 transduced from the Keio collection (114). The *rpoB8*, *rpoB2*, and *rpoB3595* mutants were P1 transduced were from the Carol Gross collection (140). Mutations were confirmed with PCR and sequencing.

Plasmids

pCA24N-Mfd is from the ASKA collection of *E. coli* ORFs without the GFP or His tag (102). Mfd is expressed from the $P_{T5\text{-}lac}$ promoter that can be activated by IPTG. pDE13 was generously provided by Dorthy Erie (University of Chapel Hill) (150).

Growth kinetics for Stl and ActD sensitivity

Overnight cultures in LB were diluted to roughly 4 x 10^6 cell/ ml in LB (containing chloramphenicol for plasmid-containing cells), and 75 µl was delivered to each well in a microtiter plate. 75 µl of Stl or ActD at twice the indicated concentration was also added to each well, for a total volume of 150 µl per well. The plate was incubated at 37°C for 12 h with constant shaking in a BioTek ELx808 Microplate Reader. The optical density (at 630 nm) of each well was read every 15 min.

For analyzing drug sensitivities, the time at which cell growth reached 50% of the maximum velocity in the absence of the drug was first calculated and considered to be the endpoint of growth. The OD at this time was then taken for each of the drug concentrations. Each drug OD was divided by the end-point OD for the no-drug control, to account for any growth deficiencies for the mutant strains. The ODs in each drug concentration was divided by the no-drug OD, then plotted against drug concentration.

Western Blots

SsrA tagging levels were analyzed in EW1b cells containing pCA24N-Mfd or the pCA24N control. Cells were pregrown overnight in LB media containing chloramphenicol at 37 °C then diluted to an OD_{560} of 0.1. The cells were grown in LB and chloramphenicol to OD_{560} 0.5 and treated with 5 μg/mL Stl or LB for 1 hour, when 4x 10^8cells were collected for each sample. Cell pellets were stored at -20 °C.

Frozen cell pellets were thawed at room temperature and resuspended in 25 μl of water and 25 μl of sample buffer (20% glycerol, 100 mM Tris pH 6.8, 2% SDS, 2% β-mercaptoethanol, and bromophenol blue), and boiled for 5 min in a water bath. An aliquot (15 μl) of each sample was loaded onto a 7.5% polyacrylamide (Tris-HCl) gel and run for approximately 2 h in 25 mM Tris-Glycine buffer containing SDS (0.1

The portion of the gel containing proteins larger than about 75 kDa was cut off and stained with Coomassie blue dye to serve as a loading control. The remaining portion of the gel was transferred to a nitrocellulose membrane for 60 min at 12 V using a Genie Blotter transfer device (Idea Scientific Co.). The blot was blocked for 1 h in 20% non-fat milk powder solution (Biorad) in Tris-buffered saline (TBS). The membrane was incubated overnight at 4°C with polyclonal *ssrA* primary antibody and Tween (0.1%), and then washed three times with TBS buffer at room temperature (10 min each). The membrane was incubated with secondary antibody (IRDye 800CW-conjugated goat anti-rabbit IgG (LI-COR®)) for 1 hour, and the washes were repeated. After air-drying, the

membrane was scanned on an Odyssey Imaging System (LI-COR Biosciences), and quantified using the Odyssey software.

***In vitro* reaction**

For *in vitro* reactions looking for full length transcript production in the presence of Mfd, open complexes were formed by adding 20 nM DNA template (254 bp PCR fragment from pDE13 containing the λ_{PR} promoter and t_{R2} terminator) and 20 nM RNAP in 1x transcription buffer (80 mM Tris-HCl pH 8.0, 200 mM KCl, 16 mM $MgCl_2$, 200 uM EDTA pH 8.0, 1 mM DTT, 0.05 mg/mL BSA and 16% Glycerol). Open complexes were incubated at 37 °C for 10 minutes. Elongation to +24 was carried out by adding NTPs (200 μM ATP, 0.3 μM ^{32}P– GTP, and 15 μM UTP) and heparin in 1X transcription buffer and incubating at room temp for 5 min. Stalled complexes were split into 25 μL each and reactions were treated with 400 μM Mfd and the indicated amount of ATP for 15 minutes at 37 °C. 200 μM of UTP, GTP, and CTP were added at room temperature for 3 minutes.

For the reaction monitoring the competition of Mfd and elongation, stalled complexes were formed as above and split into 25 μL reactions. For the reactions labeled "Mfd added first, 400 μM Mfd was preincubated with stalled complexes for 12 minutes at the indicated temperature. All four nucleotides (200 μM each) were then added for 3 minutes. For the reactions labeled "all added together", 400 μM Mfd was

added with all four nucleotides (200 μM each) and incubated for at the indicated temperature and time.

For the ActD dilution assay, ActD stalled complexes were treated with 200 μM of UTP, CTP, and GTP, 4 mM ATP, and 400 nM Mfd (or storage buffer) and incubated at room temp for 10 minutes. 1x transcription buffer was added to dilute the reaction 4-fold, resulting in an ActD concentration of 0.2 μM, and incubated at room temp for 3 min.

All reactions were quenched with 50% formaldehyde and 10x native loading dye and placed on ice until being loaded in the gel. Samples were loaded into 8 M urea, 20% polyacrylamide native gels and run at 100 V until the loading dye had run about 9/10 of the gel length. The bottom portion of the gel (containing the majority of the free nucleotides) was excised and the gel was exposed to a phosphoimager screen overnight.

5. Mfd alters global transcription patterns in undamaged *Escherichia coli* cells

5.1 Introduction

Mfd, the transcription-coupled repair factor in *E. coli*, is a well-studied enzyme known for decreasing mutation frequency and facilitating strand-specific DNA repair (66,91,151,152). In addition, recent literature has expanded the known roles of Mfd to reach beyond transcription-coupled repair in *E. coli*. As shown in Chapter 3, cells lacking Mfd are resistant to Streptolydigin, implying a role for Mfd in termination of stalled transcription complexes in the absence of any DNA damage (84). Mfd has been shown to terminate transcription stalled by nucleotide deprivation (also without any DNA damage) (147). Selby and Sancar found evidence that suggests Mfd might act on paused RNAP (69). It has also been suggested by Hanawalt and colleagues that transcription paused at natural pause sites are subject to gratuitous TCR (153), though it has never been directly proven. Accordingly, it stands to reason that Mfd might play a role in transcription termination in undamaged cells. The purpose of this Chapter is to identify the effects of Mfd on global transcription patterns in undamaged cells.

Mechanistically, Mfd is recruited to a stalled elongation complex through its RNAP binding domain D4, binding directly upstream of the transcription bubble (30,67). Subsequent DNA binding leads to a conformational shift, causing movement of the helical hairpin into the dsDNA binding cleft. This hairpin serves as a ratchet, pushing RNAP forward and DNA backwards. Depending on various factors such as sequence

and RNAP conformation, Mfd then either enhances forward translocation or removes the stalled RNAP from the DNA. For instance, if RNAP is backtracked, Mfd uses its ATP dependent translocase activity to push forward backtracked RNAP. This concomitantly rewinds the transcription bubble from upstream and unwinds the RNA/DNA hybrid, allowing for resumption of elongation (67,68). When there is a physical obstruction in the way, such as a DNA-bound protein or DNA lesion, this ratcheting motion is proposed to dislodge RNAP from DNA (66,69). If the block was induced by DNA damage, once RNAP has been cleared from the DNA UvrA is recruited to the site via Mfd's UvrA binding site in the N-terminal UvrB homology domains (66,67), initating nucleotide excision repair. If there is no DNA-damage, such as the case of natural pauses, Mfd is potentially capable of terminating transcription in a manner that is dependent on the conformation of RNAP.

There are three types of pauses to consider in *E. coli*: class 1 (hairpin dependent pauses), class II (backtracking), and ubiquitous pauses. Class I pauses inhibit elongation through RNA hairpin–RNAP interactions that stabilize the RNA 3′ end in a frayed or hypertranslocated position (154,155). Backtracking in class II hybrids is caused by weak RNA:DNA hybrids; backtracking causes upstream sliding of RNAP and occlusion of the active site with nascent RNA (156-158). Ubiquitous pauses are less well studied, but are proposed to involve structural rearrangements of RNAP that result in a temporarily inactivated intermediate (136). This inactivated intermediate can either be restarted

(resulting in the appearance of a ubiquitous pause) or result in a longer-lived pause (class I or class II) (33). Single molecule studies show that ubiquitous pauses do not backtrack (33).

For Mfd to be able to terminate transcription at pause sites or other locations, certain requirements must be met. First, RNAP must already be stalled; it has been shown that Mfd has no effect on either Rho-dependent or independent terminators (69), and Mfd is outcompeted by elongation at physiological levels of NTPs (147). Second, the σ factor must be released, as Mfd requires 25 bases of DNA template upstream of the stalled RNAP to act on stalled RNAP, and this site is occupied by σ when present (9,147). It has also been shown directly that the presence of σ^{70} blocks Mfd's ability to remove stalled polymerases (147). Furthermore, the RNAP would most likely need to be stalled in such a way that it was not backtracked, as this would simply result in forward translocation (see above).

In this Chapter, we use RNAseq technology as well as cell-based assays to further examine the role of Mfd as a transcription regulator in undamaged cells. We show evidence that alterations to Mfd levels in the cell cause vast global transcriptome changes, and suggest that Mfd might regulate the transcription of genes that affect growth rate and DNA replication. We also investigate the location of Mfd-induced termination, with a particular focus on the possibility of termination occurring at pause sites, and find that surprisingly Mfd appears to be affecting gene transcription at or near

promoters. Taken together, this further expands the role of Mfd well beyond that of a simple repair factor.

5.2 Results

Mfd has termination activity in wildtype cells

While *mfd* cells have no noticeable phenotype during normal growth (84), we have previously noticed that cells containing plasmids that carried the Mfd gene have growth deficiencies. EW1b cells containing pMfd19, which overexpresses Mfd from its natural promoter, grew much more poorly than EW1b cells containing a vector plasmid: the cells have a decreased maximum growth rate and a lower maximum cell density when grown in a liquid culture (data not shown and (84)). Overexpressing Mfd from the IPTG inducible pCA24N-Mfd plasmid, which expresses Mfd at high levels, is lethal when grown in our standard sensitivity assay (data not shown). This implies that too much Mfd activity in wildtype cells is detrimental to cell growth.

Given our previous findings with Mfd (See Chapter 4), we wanted to see if the detrimental activity of Mfd was premature termination. If Mfd is causing premature termination, it would be creating substrates for the tmRNA system. As shown in Figure 23, we found evidence for tmRNA tagging in Mfd-overexpressing cells even in the absence of Stl. To further characterize this tagging, we performed western blots using an antibody against the degradation tag of tmRNA and looked for the presence of

tmRNA-terminated peptides during various times of Mfd overexpression. We found that tmRNA tagging was indeed induced in a time dependent manner during overexpression of Mfd, with detectable peptide accumulation at about 30 minutes (Figure 26). This data implies that Mfd has low levels of termination activity in WT cells, and this activity is exacerbated when Mfd is overexpressed. One possibility is that Mfd is terminating elongation in undamaged cells at natural pause sites, which in some conformations might be fairly similar to Stl-stalled polymerases.

Testing for DNA damage due to aberrant Mfd activity

It has been suggested that RNA polymerase stalled at a natural pause site might lead to gratuitous TCR, leading to DNA damage and possible mutations (153). While never directly proven, there have been low levels of Uvr-ABC excision detected on undamaged DNA *in vivo* and *in vitro* (159). Therefore aberrant Mfd activity due to overexpression might lead to recruitment of UvrA to undamaged sites. We tested *uvrA* knockouts to see if deleting *uvrA* alleviates cell sensitivity to pMFD19. We found no difference in cell growth in EW1b *uvrA* pMfd19 cells compared to EW1b pMFD19 cells (Figure 27A). We also treated both cell lines with Streptolydigin to stimulate stalling and found no difference (Figure 27B). Therefore we were unable to detect UvrA-dependent gratuitous TCR.

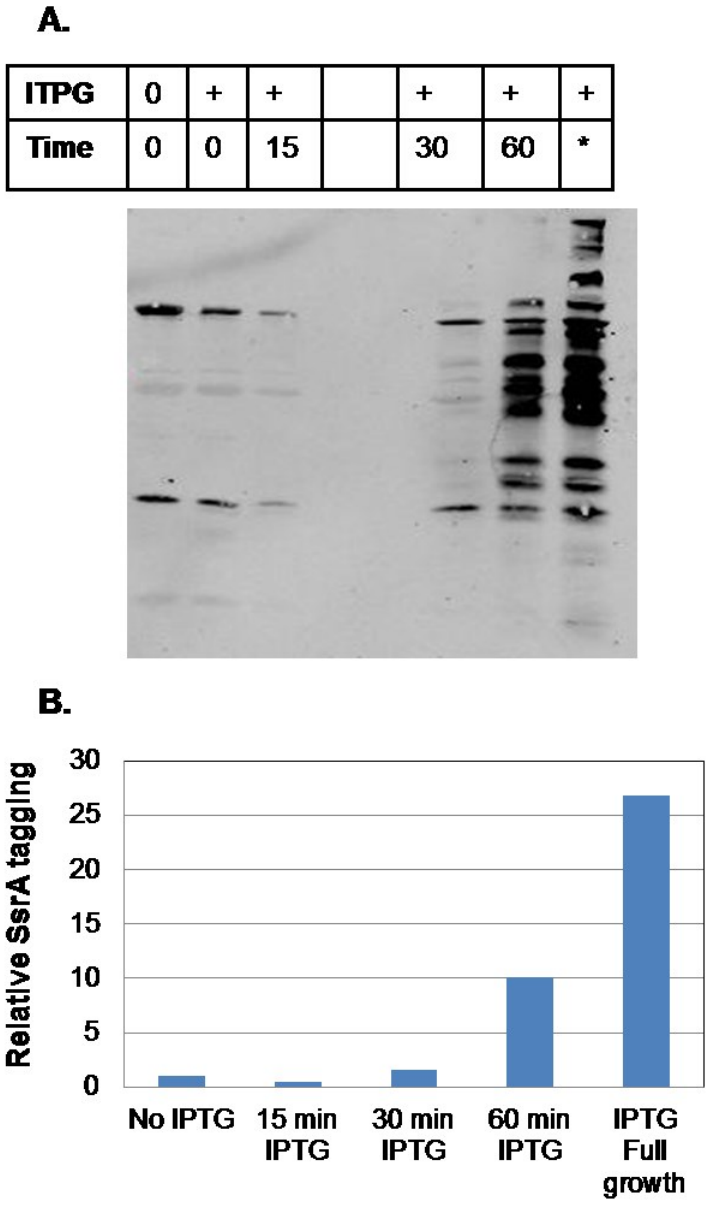


Figure 26: tmRNA tagging in cells overexpressing Mfd.
Panel A: Extracts from MG1655 cells containing pCA24N-Mfd were analyzed by Western blotting with polyclonal antibodies to the degradation tag of tmRNA. Western blot was conducted as stated in Materials and Methods. Panel B: Tagging levels are expressed as a ratio of tmRNA tagging levels relative to the no drug control.

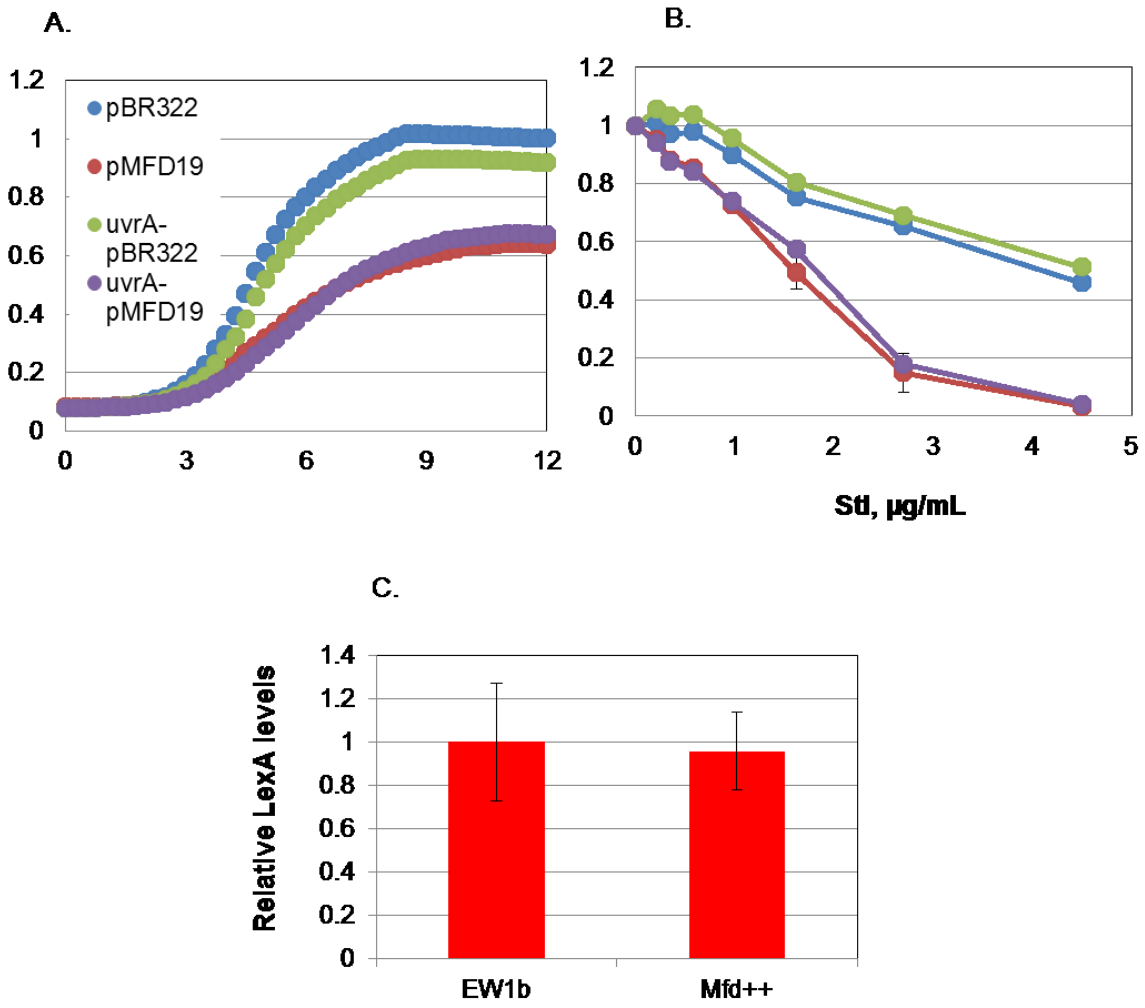


Figure 27: Sensitivity of *uvrA* cells to expression of Mfd.
EW1b and EW1b *uvrA* cells containing either pMfd19 or the parent pBR322 plasmid were tested for sensitivity to Mfd expression and transcription elongation inhibitor Streptolydigin using a microtiter plate sensitivity assay. Panel A is representative growth curves for the wild type and mutant cell lines respectively without any Stl treatment. Panel B is the comparative titration curve using 50% of the maximum exponential rate as an endpoint. Panel C is the quantitation of the LexA western blots. Western blots and analysis were performed as in Figure 25 except polyclonal LexA antibodies were used as the primary antibody. Error bars are representative of the standard deviation of three independent cultures.

We also wanted to see if overexpressing Mfd led to UvrA-independent DNA damage, and therefore tested for SOS induction after Mfd overexpression. We attempted to perform quantitative RecA western blots as a measure of SOS, but we found that RecA levels were greatly decreased in Mfd overexpressing cells (data not shown). This is presumably due to Mfd affecting *recA* RNA levels, most likely by terminating its transcription. We then chose to perform Western blots using a primary antibody against LexA, as monitoring LexA degradation is another accepted way to look at SOS induction and would not be affected by alterations to transcription (81). We found the LexA levels were unchanged in Mfd-overexpressing cells, indicating a lack of DNA damage in these cells (Figure 27C).

Mfd alters transcription patterns in undamaged cells

Since Mfd-overexpression appears to be lead to termination in WT cells, it reasonably follows that Mfd might be affecting global transcription patterns. Therefore, we performed RNAseq analysis on three MG1655 *mfd* cells lines and looked for expression changes when compared to MG1655 cell lines. We chose MG1655 as our background as it has been fully sequenced and annotated and is commonly accepted as a standard K-12 strain (160). Using three independent control and mutant cells lines for comparison allowed us to normalize the data using the TMM method, which adjusts for variations in RNA levels between isolates. We can then calculate statistically significant

differences between cell lines, and therefore all genes included in further analysis were those that passed a significance cut-off p>0.05.

Under the conditions listed above we found that 94 genes were affected 2-fold or more in the *mfd* cells, representing 2% of annotated genes in MG1655. Of these, a total of 75 genes are overexpressed more than 2-fold in *mfd* cells, and an additional 116 are overexpressed 1.5-fold to 2 fold in *mfd* cells (Table 6). Furthermore, we found 18 genes were repressed more than 2-fold in *mfd* cells (along with *mfd* itself), and another 17 genes were repressed 1.5-2 fold (Table 7). Given the uneven distribution of overexpressed versus repressed genes (75 vs 18), it appears as if Mfd preferentially acts as a negative regulator of gene expression.

To categorize these genes we utilized the Database for Annotation, Visualization and Integrated Discovery (DAVID) program to sort genes based on their GO-terms (161,162). P-values for significance were calculated using the Fisher Exact test. Functional analysis using GO-terms of the overexpressed genes showed that ncRNA metabolic process genes, nucleobase, nucleoside and nucleotide interconversion genes, and DNA replication genes were well represented (Table 8A). Nearly half of the repressed genes are involved in carbohydrate catabolic processes (Table 8B). When the genes that are overexpressed 1.5-fold to 2-fold are included, roughly 10% (20 out of 191) of the total genes are ncRNA metabolic process genes, giving a Fisher Exact p-value of

Table 6: Genes overexpressed >2-fold (p>0.05) in *mfd* cells

Upregulated genes	Fold difference	Upregulated genes	Fold difference	Upregulated genes	Fold difference
yhfY	9.82	*mscS*	2.50	*aroH*	2.20
ypdK	4.43	*rzoR*	2.46	*valZ*	2.18
ynfA	3.84	*iraM*	2.43	*dnaG*	2.17
yghG	3.50	*appY*	2.42	*yfcJ*	2.17
yghF	3.26	*gltD*	2.39	*yecF*	2.16
nrdF	3.17	*rsxA*	2.37	*rnc*	2.16
yliF	3.16	*ybjG*	2.36	*pabA*	2.16
yceA	3.14	*tadA*	2.35	*valT*	2.14
omrA	3.14	*phoU*	2.34	*ykiA*	2.13
lysT	3.12	*leuW*	2.34	*yojI*	2.12
rzpD	3.04	*mltF*	2.34	*udk*	2.12
apt	2.99	*yjeA*	2.34	*ydgK*	2.11
pyrF	2.84	*lysW*	2.33	*yadS*	2.11
pppA	2.79	*mgrR*	2.32	*ygdQ*	2.11
ycaD	2.78	*ybjX*	2.32	*rlmD*	2.10
yliE	2.76	*rfaH*	2.30	*rsmG*	2.10
queD	2.76	*dnaA*	2.30	*yjhD*	2.09
lysQ	2.71	*ybgT*	2.29	*rzpR*	2.08
yciH	2.64	*rsxB*	2.27	*dusC*	2.08
lysY	2.64	*mnmG*	2.26	*gltB*	2.07
lysZ	2.63	*ybgE*	2.26	*slyX*	2.06
rlmF	2.59	*lysV*	2.25	*dnaN*	2.05
gpt	2.59	*tusA*	2.24	*ileU*	2.04
gltW	2.58	*cydB*	2.24	*secE*	2.03
omrB	2.58	*yghE*	2.23	*yhfZ*	2.02

Table 7: Genes repressed >2-fold (p>0.05) in *mfd* cells

Downregulated Genes	Fold Difference
mfd	14.03
nanA	4.86
malF	4.19
malK	3.98
astC	3.62
yihN	3.34
malM	3.34
yihM	3.07
fruB	2.77
dgsA	2.65
fumC	2.54
fruK	2.33
malS	2.20
rbsD	2.20
sthA	2.19
sdhD	2.13
malI	2.10
sdhA	2.07
clpS	2.06

Table 8a: GO annotations with significant representation in genes overexpressed >2-fold in *mfd* cells.

GO Term	Role	Genes	%	p-value
GO:0015949	Nucleobase, nucleoside and nucleotide interconversion	*pyrF, apt, nrdF, gpt, udk*	6.7	3.70E-07
GO:0034660	ncRNA metabolic process	*tusA, tadA, rsmG, rnc, mnmG, rlmF, yjeA, dusC*	10.7	3.10E-06
GO:0006261	DNA-dependent DNA replication	*dnaA, dnaG, dnaN*	4	5.20E-03
GO:0044271	Nitrogen compound biosynthetic process	*pryF, aroH, apt, gltD, gpt, pabA, queD, udk, gltB*	12	1.70E-03
GO:0008219	Cell death	*rzpR, rzpD, rzoR*	4	1.00E-02

Table 8b: GO annotations with significant representation in genes repressed >2-fold in *mfd* cells

GO Term	Role Category	Genes	%	p-value
GO:0016052	carbohydrate catabolic process	*malM, fruK, malK, malI, malF, fruB, malS rbsD*	44.4	3.50E-08
GO:0046356	acetyl-CoA catabolic process	*sdhA, sdhD, fumC*	16.7	6.80E-05

1.2×10^{-12}. In addition, the number of genes involved in DNA replication increases to 6% (11 out of 191) with a p-value of 4.1×10^{-9}.

One indirect model for the effects on transcription is that Mfd could potentially only be affecting transcription of a one or a few transcription factors, which would then have a more global downstream effect. Therefore we searched for known transcription factors as listed on RegulonDB among our list of affected genes and cross-referenced the lists of overexpressed and repressed genes with the genes that those transcription factors regulate (163). There were two overexpressed and two repressed transcription factors in *mfd* cells, namely *appY* and *dnaA* (overexpressed) and *dgsA* and *malI* (repressed). However, none of the genes that these transcription factors regulate were changed in *mfd* cells (data not shown). Therefore it appears as if Mfd is having a direct effect on gene expression.

Mfd potentially targets transcription of RNAs

In addition to the GO annotations noted previously, which only consider protein coding genes, we noticed that several of the genes that are overexpressed in *mfd* cells are tRNA genes. In fact, 11 of the 75 genes overexpressed >2-fold are tRNA molecules: *lysT, lysQ, lysY, lysZ, leuW, lysW, lysV, valT, ileU, valZ, gltW*. An additional 8 are upregulated 1.5-2 fold. Of the 11 tRNA genes found in ribosomal operons 7 were found to be overexpressed >1.5-fold, indicating that Mfd potentially targets rRNA operons.

Unfortunately, we cannot monitor rRNA levels using RNAseq, as the high expression of rRNA compared to tRNA and mRNA require them to be removed during library preparations. 10 of the 19 tRNA genes are found on two clusters of tRNA genes outside of the rRNA operons, suggesting that these clusters might be operons also regulated by Mfd. In addition to the tRNA 3 other genes overexpressed >2-fold (*omrA, omrB, and mgrR*) are small regulatory RNAs, based on the list of small regulatory RNA identified in the EcoCyc database (164) (p-value 0.035). Taken together with the large number of RNA modifiers overexpressed in *mfd* cells, this data strongly implies Mfd plays a role in negatively regulating RNA biology.

Potential connection to growth rate regulation

The specific targeting of tRNA genes, particularly those in ribosomal RNA operons, RNA modifiers, and DNA replication also suggests Mfd might play a role in growth rate regulation. Growth rate determination in *E. coli* is primarily controlled by the number of ribosomes available in the cell, with the rate of rRNA synthesis being most sensitive to changes in growth conditions (165-168). rRNA synthesis is control by the two alarmone molecules ppGpp and pppGpp, affecting both stringent response and growth rate regulation (169). The transcription factor DksA also acts synergistically with ppGpp, although DksA also plays a ppGpp independent role on gene regulation (170,171). Therefore we compared the genes affected by *dksA* and ppGpp deficiencies to

those affected by a *mfd* deficiency. We found that of the 18 genes downgregulated in *mfd* mutants, 6 genes were also downregulated in *dksA* mutants (*fruK, fumC, malF, malK, malM, and sdhA*), giving a p-value of 6.21×10^{-6} using a Fisher exact test. Four of those genes (*malF, malK, malM, and fumC*) were also downregulated in the ppGpp^{o} strain, giving a p-value of 1.27×10^{-3}. Of the genes overexpressed in *mfd* mutants, only four genes (*apt, pyrF, yciH, and yecF*) were overexpressed in a *dksA* strain (p-value 0.29). It is of note that the experiments with the *dksA* and ppGpp^{o} strains excluded all rRNA and tRNAs, and therefore they were excluded from this comparison. However, ppGpp and DksA stimulate expression of rRNAs, while Mfd apparently represses transcription of tRNAs, especially those included in rRNA operons. Taken together, it appears as if Mfd may contribute to growth rate regulation, but in a manner that only partially overlaps the function of ppGpp and DksA. See Discussion for further comments.

Mfd-overexpression leads to dramatic global transcription pattern changes

Given that overexpressing Mfd lead to tmRNA tagging, we also tested for changes in global gene expression in Mfd-overproducing cells. Two MG1655 pCA24N-Mfd cell lines were submitted for RNAseq analysis, along with two MG1655 pCA24N cell lines for controls, using the same growth and sequencing conditions as previously stated. ANOVA analysis was used to compare results from the two groups, with the caveat that using only two cell lines for each does not allow us the same statistical

confidence as before. Mfd overexpressing cells had a dramatic change in global transcription patterns, with 103 genes overexpressed more than 2-fold (including Mfd) and 1019 genes repressed 2-fold or more (data not shown). However, when the list of genes that are overexpressed in *mfd* cells are cross referenced to the genes that are repressed in Mfd-overexpressing cells, only 10 genes overlap (out of a potential 75). Therefore it appears as if Mfd has different targets when Mfd levels are artificially high.

One trivial reason for the vast genomic changes in Mfd overexpressing cells could be due to generic changes due to protein overexpression, rather than specific activity of Mfd. Therefore we compared the list of genes overexpressed and repressed 2-fold in Mfd overexpressing genes to the list of genes found to be affected by recombinant protein overexpression in MG1655 pPROEx-CAT containing cells, which overexpresses chloramphenicol acetyltransferase on an IPTG inducible promoter (172). Surprisingly, only 71 of the repressed genes and 2 of the overexpressed genes in Mfd-overexpressing cell lines overlapped with the genes overexpressed and repressed, respectively, in the MG1655 pPROEx-CAT cells (data not shown). While our data does not quantitatively list the degree to which genes are affected by increases in Mfd activity, it is apparent that Mfd activity does induce gross genomic changes that are not simply due to generic recombinant protein affects.

Mfd-overexpressing cells have preferential transcription changes in the terminus region

To see if Mfd is preferentially targeting genes in a particular area of the chromosome we plotted the location of each of the genes that are overexpressed and repressed in *mfd* cells. When plotted along chromosome position the affected genes seem to be fairly evenly distributed (Figure 28A). Genes that are overexpressed in Mfd-overexpressing cells also were fairly evenly distributed (Figure 28B). In contrast, read coverage profiles of for each of the Mfd-overexpressing cell lines show that transcription of the majority of the terminus region is downregulated (Figure 29). When the gene RPKM values for Mfd overexpressing cells are compared to WT cells, it is further apparent that there is a cluster of genes in the terminus region that is downregulated (Figure 28C). Specifically, 92 of the 250 annotated genes between TerA and TerC are downregulated (Fisher p-value = 0.00007). One possibly is that overexpressing Mfd causes downstream structural rearrangements of the terminus region (See Discussion).

Gene expression deviates at or near promoters

The use of RNAseq allows us to look specifically at where transcription is deviating between WT and *mfd* cell lines. We have hypothesized that Mfd is terminating transcription during elongation, possibly during natural pauses of RNA polymerase. If this is true, one might expect to have a higher chance to find specific sites of termination

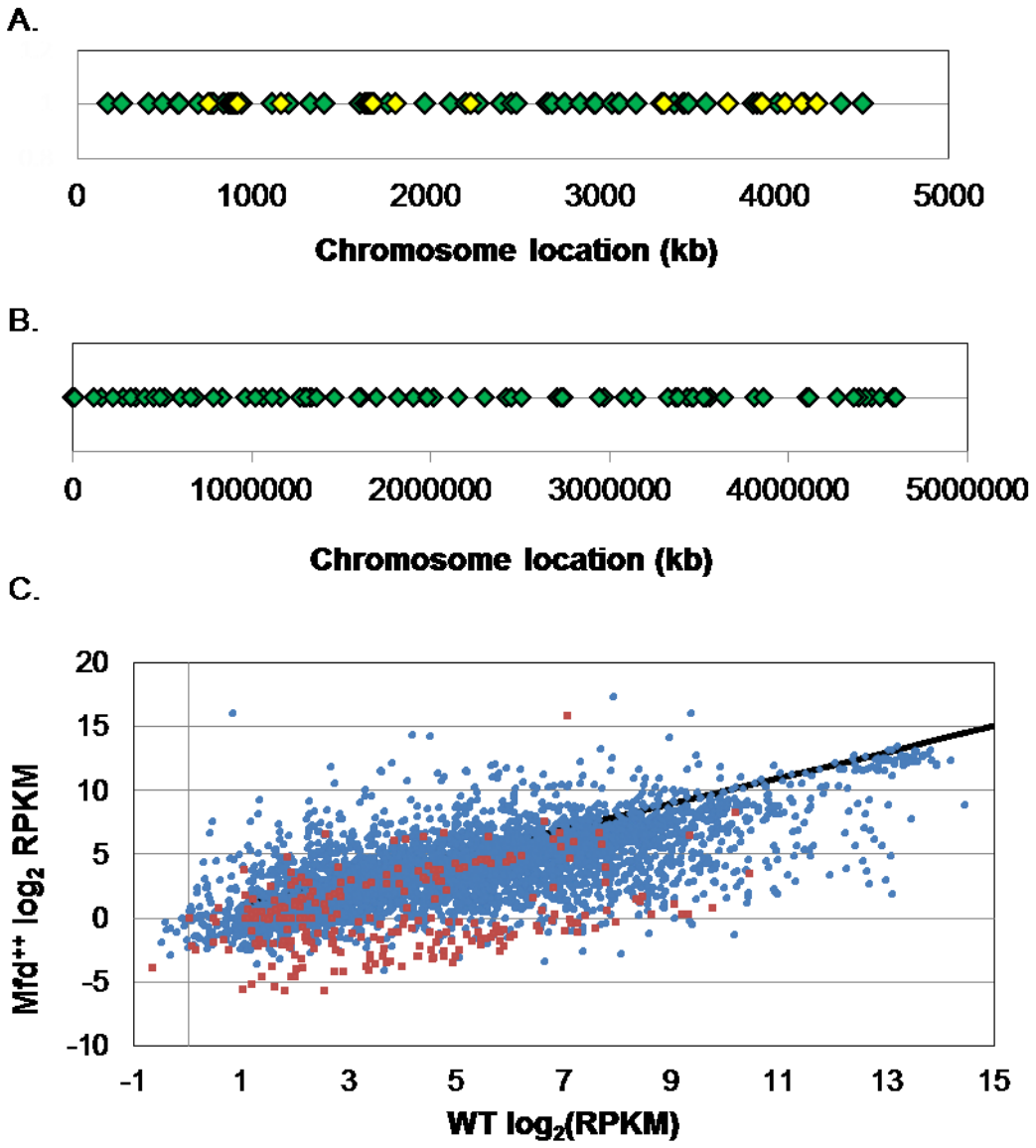


Figure 28: Location of overexpressed and repressed genes in *mfd* cells.
Panel A is the chromosome position of overexpressed (green diamonds) and repressed (yellow diamonds) genes in *mfd* cells. Panel B is the chromosome position of overexpressed (green diamonds) genes in Mfd overexpressing cells. Panel C is the MG1655 pCA24N $\log_2$ RPKM values plotted against MG1655 pCA24N-Mfd $\log_2$ values. A perfect 1:1 ratio is plotted as a black line. Genes falling off this line are under- or over-represented in Mfd-overexpressing cells. Genes in the terminus region (between TerA and TerC) are plotted in red. Genes in the terminus region that are repressed in Mfd-overexpressing cells are boxed in red.

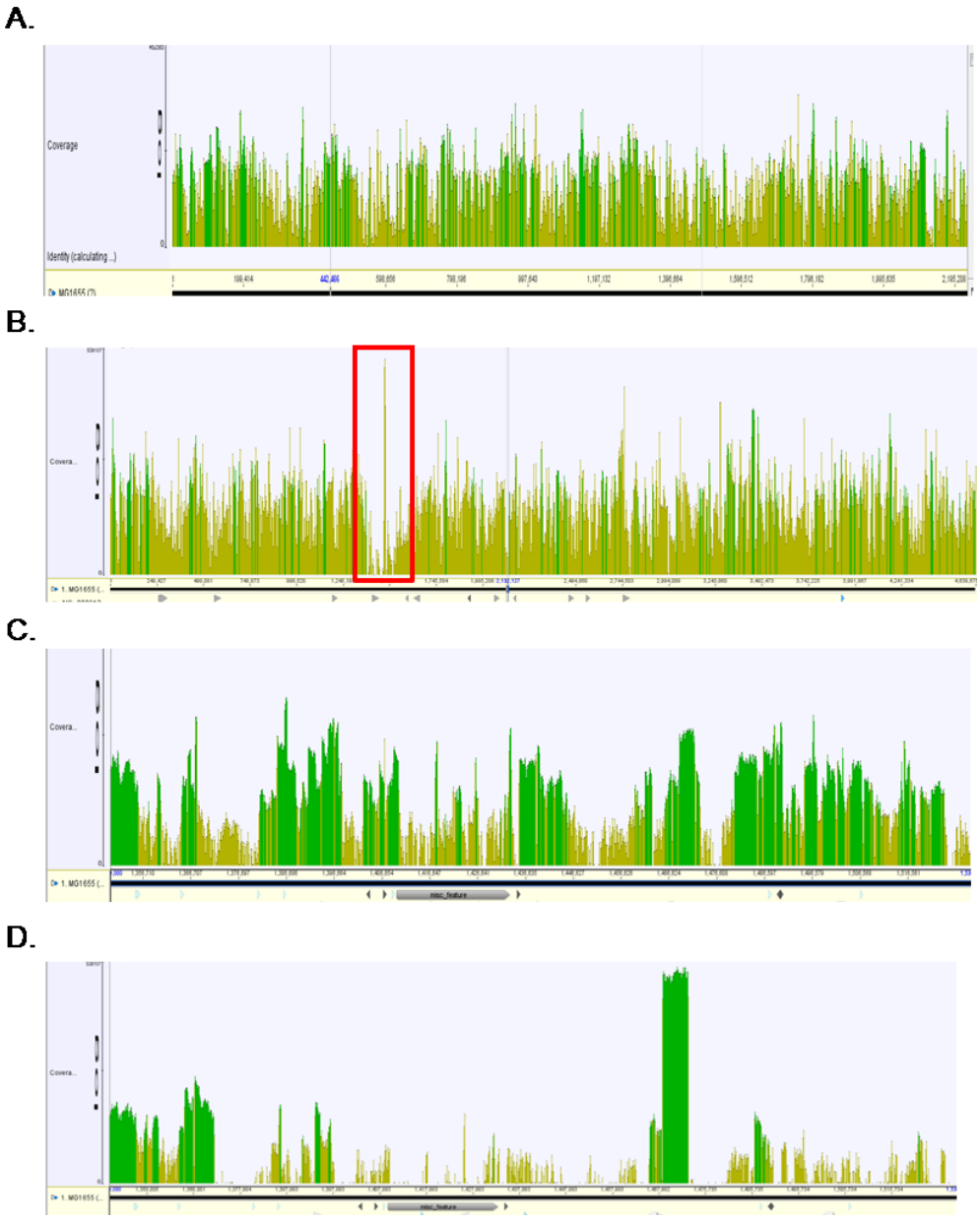


Figure 29: Read coverage profiles for WT and Mfd-overexpressing cells
Panels A and B are the read coverage profiles generated by Geneious Pro for the entire chromosome for a representative MG1655 (A) and MG1655 pCA24N-Mfd (B) cell line. The highly underrepresented Ter region in MG1655 pCA24N-Mfd cells is highlighted (red box). Panel C is the read coverage profiles generated by for a representative MG1655 pCA24N cells in the terminus region (between TerA and TerC). Panel D is the read coverage profiles for a representative MG1655 pCA24N-Mfd cell line in the same terminus region as C.

in longer operons. Therefore, we identified all the genes that are overexpressed 2-fold or more in *mfd* knockout cells that belong to operons. A total of 22 operons included overexpressed genes. Using JMP Genomics and Microsoft Excel, read coverage maps were generated for both WT and *mfd* strains by averaging the read coverage per every 100 bp, then averaging the three replicates. Read coverage for both cell types was then plotted as a function of chromosome position. Surprisingly, read coverage for all deviated from WT at or near the promoter for all operons examined (see Figure 30A). We did not find any evidence for decreases in read coverage in the wildtype cells compared to the *mfd* knockout read coverage within the body of the operon, as would be expected if Mfd is terminating transcription in wildtype cells.

To get a more precise idea of where WT and *mfd* gene expression were deviating, average read coverage for every 5 base pairs was calculated and plotted for *dnaN*, an example gene that is overexpressed in *mfd* cells. While the pattern of read coverage was consistent between the two strains, it was clearly apparent that the level of RNA in *mfd* cells deviated from wildtype near the promoter (Figure 31A). To pinpoint the exact location of deviation, the ratio of read coverage was taken for *mfd*/WT cells and plotted, with the deviations from a ratio of 1 indicating the location where transcription was affected. Transcription deviated <20 bp from the promoter (Figure 31B). This data contradicts our model that Mfd is terminating during elongation. It instead indicates

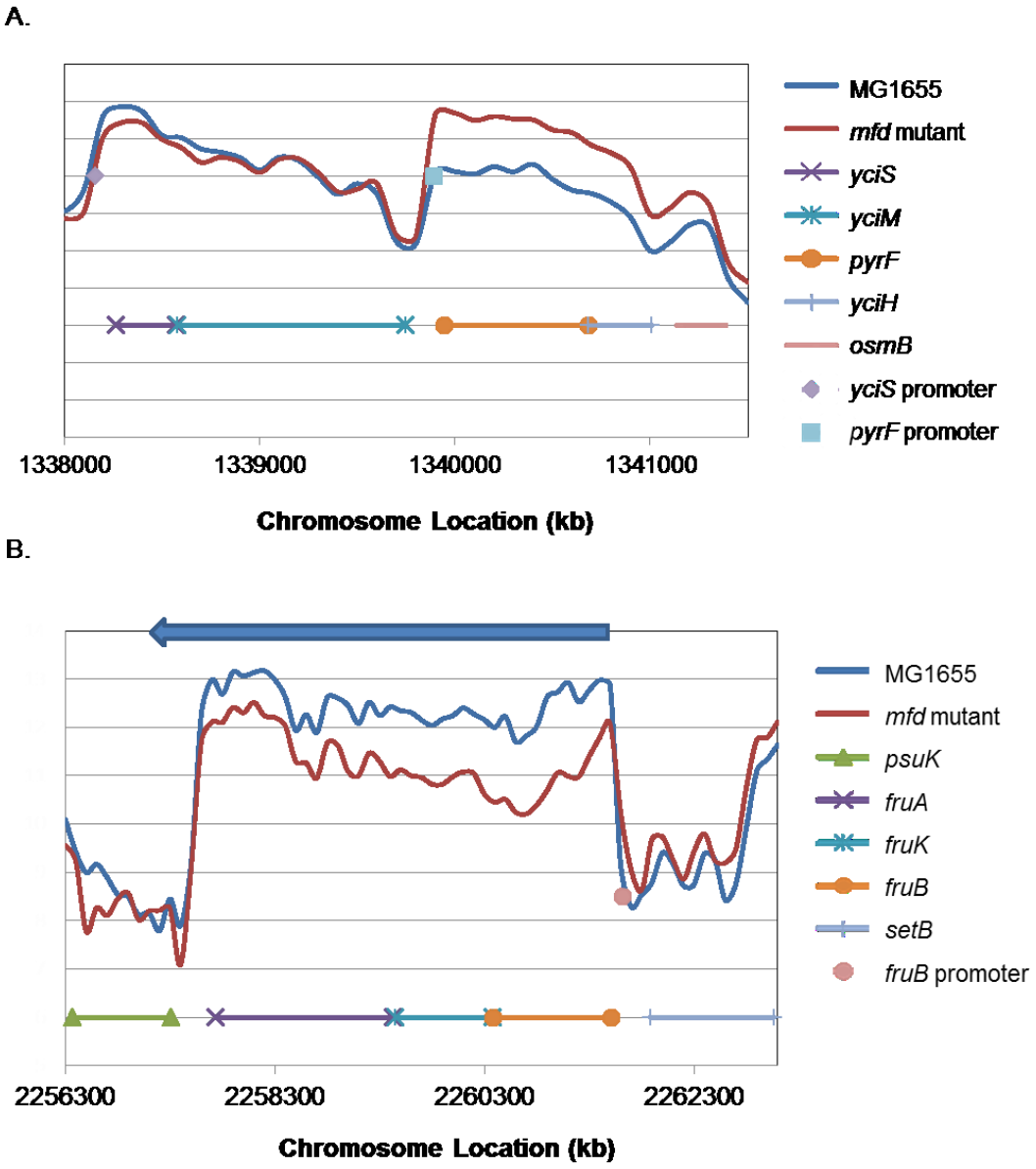


Figure 30: Representative read coverage profile for MG1655 and *mfd* cell lines
Panel A is the read coverage profile of a representative operon including genes that are overexpressed in *mfd* cells. Curves were generated by connecting the $\log_2$ RPKM data for 100 bp bins. The four genes, *yciS*, *yciM*, *pyrF*, and *yciH* are all part of one operon, with a second promoter present which transcribes *pyrF* and *yciH* separately. Transcription proceeds from left to right. Panel B is a representative operon including genes that are repressed in *mfd* cells. *fruB*, *fruK*, and *fruA* are expressed together starting from the *fruB* promoter. Transcription proceeds from right to left (blue arrow). Operon information was taken from Ecocyc (164).

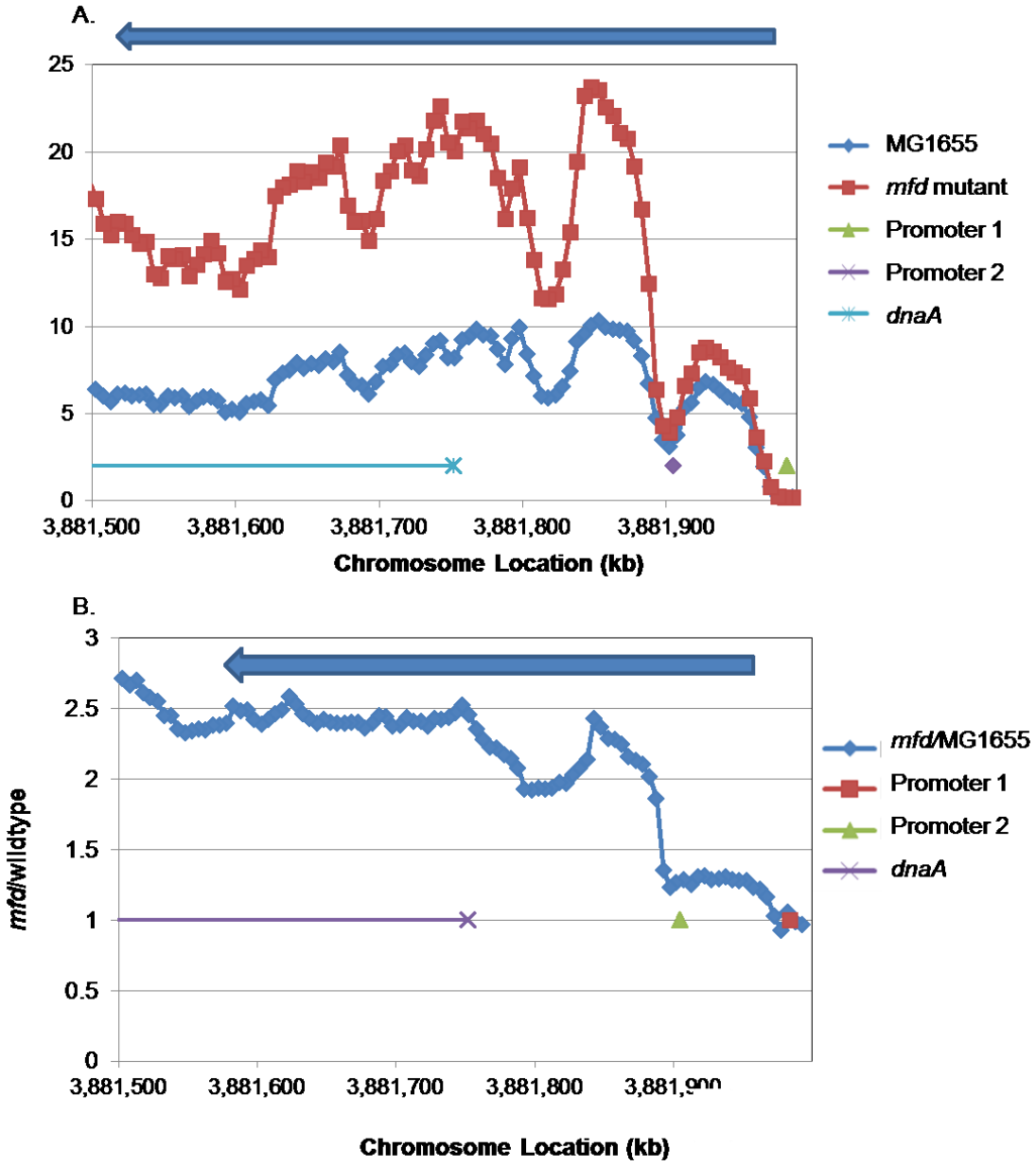


Figure 31: Representative read coverage profile for MG1655 and *mfd* cell lines using 5 bp bins.

Panel A is the read coverage profile for the *dnaA* gene generated from the $\log_2$ RPKM values for 5 bp bins. Transcription proceeds from right to left (blue arrow). Read coverage from *mfd* cell line (red) deviates from MG1655 coverage (blue) within 15 bp from each promoter (green triangle and purple diamond). Panel B is the ratio of *mfd* RPKM values to wildtype RPKM values plotted against chromosome location. Deviations from 1 represent overexpression or repression.

that Mfd is either directly terminating transcription at the promoter, or causing a secondary effect at the promoters of these genes.

Given that Mfd does not appear to be terminating during elongation, we were then curious about the reverse condition, where Mfd activity appears to be promoting gene expression (represented by the downregulated genes in the *mfd* mutant). If Mfd is negatively affecting gene production near promoters, is it also capable of enhancing gene production at promoters? We found 8 operons included repressed genes, and plotted the read coverage maps using the 100 bp bins RPKM data. Like the overexpressed genes, *mfd* coverage deviated from the wildtype coverage near promoters (Figure 30B). Analysis using 5 bp bins of the *malK* gene (which is repressed in *mfd* cells) confirmed the deviation occurred <20 bp from the promoter (data not shown).

To broaden our search for termination sites outside of operons we looked for locations in the gene coverage maps where there is a sharp drop-off in read coverage in the MG1655 coverage maps that were not present in the *mfd* coverage. This would indicate a position where Mfd is terminating transcription in wildtype cells but not in *mfd* cells. Therefore we used the RPKM data for the 100 bp bins, and averaged 3 bins together, starting with every bin, to smooth out the noise in the data. The slope in-between each averaged bin set was taken for both cell lines and plotted against each other. All *mfd* slopes and WT slopes less than -1 (as we were looking for sharp drops in

WT cells) were plotted against each other to highlight outliers with negative WT slopes and neutral or positive Mfd slopes (see Figure 32A). Each transition examined occurred in-between genes (see Figure 32B for a representative example), further emphasizing that Mfd effects were occurring near promoters. As a control, we examined the reverse conditions, where there were sharp drops in *mfd* cell lines that were not in the MG1655 cells lines. The only dramatic read changes were due to the sharp drop in read expression at the *mfd* gene in the *mfd* cell lines as expected, confirming the validity in our method in identifying *bona fide* expression level transitions.

As a large number of genes were downregulated in Mfd-overexpressing cell lines, we also conducted a similar analysis of read coverage in MG1655 pCA24N cells vs MG165 pCA24N-Mfd cell lines. Not surprisingly, we found numerous points where there was a large drop in Mfd-overexpressing cells that was not present in WT cells. After examining several of these points, it became quickly apparent that the large number of outliers were due to transitions from repressed genes to genes with similar expression as WT. Therefore, we narrowed our search to look for transitions that occurred within genes. Using both bin RPKM data and gene annotations , we screened through the genes with more than a 5-fold change in expression from beginning to end (roughly 20 genes) and identified two genes with the desired characteristic in both Mfd-overexpressing strains: *ycjM* and *macB* (See Figure 33). Structural analysis of the sequence surrounding the drop did not suggest the presence of an RNA hairpin,

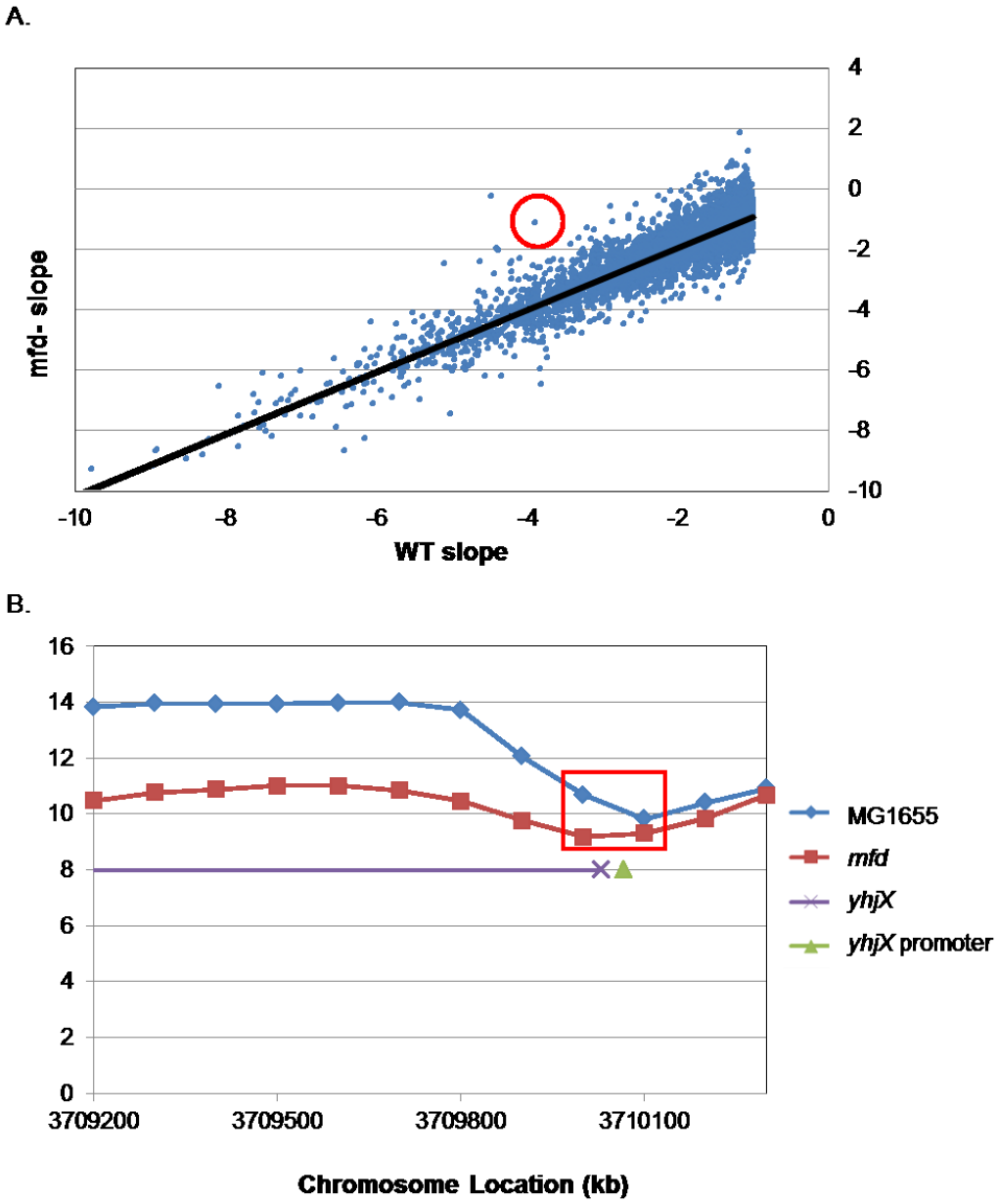
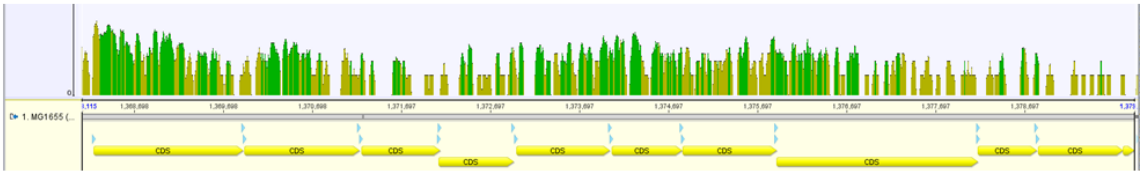


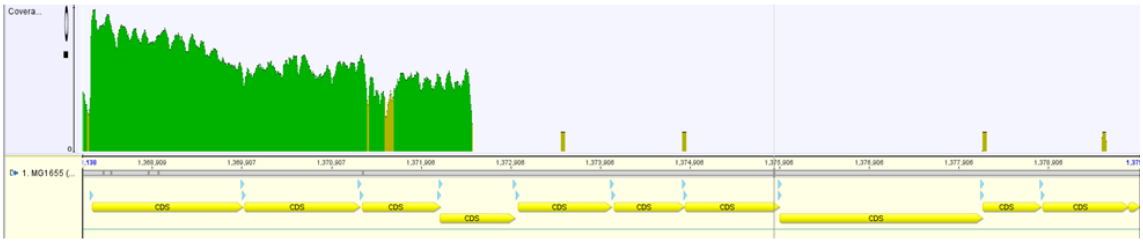
Figure 32: Sharp drops in read coverage in WT cells

Panel A is the comparison of changes in read coverage for each 100 bp averaged bin for WT (MG1655) and *mfd* cells. Locations of interest are those with a large negative WT slope and a small *mfd* slope. The point circled in red is the location examined in Panel B. Panel B is the read coverage profiles for MG1655 and *mfd* cells, with the transition of interest shown in the red box. While this is the target profile, it does not indicate a drop in transcription coverage within a gene, and instead reflects the transition in-between genes.

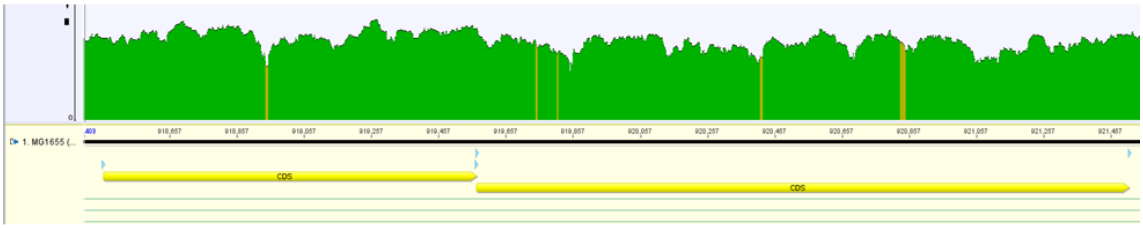
A. MG1655 pCA2N



B. MG1655 pCA24N-Mfd



C. MG1655 pCA2N



D. MG1655 pCA24N-Mfd

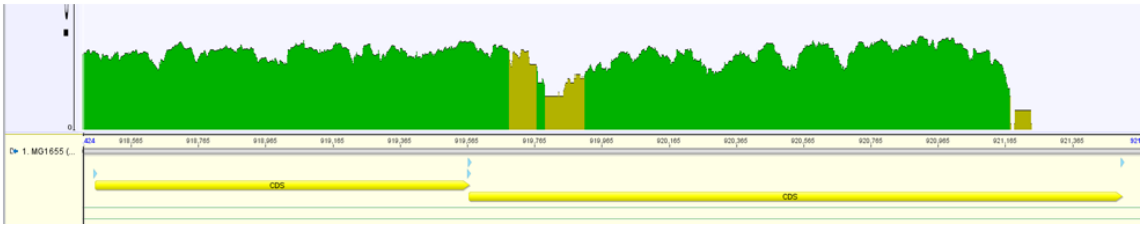


Figure 33: Sharp coverage drops within genes in Mfd-overexpressing cells.
Read coverage profiles for the genes with a sharp cut-off in the middle of the gene generated by Geneious. The drop off in Panel B occurs in the *ycjM* gene, and the drop off in Panel D occurs in the *macB* gene.

strongly suggesting against a possible type 1 pause site as the cause of elongating pausing/termination.

To further examine the possibility of pause sites as Mfd termination sites, we examined the read coverage profiles for known type 1 (hairpin dependent) and type 2 (backtracking) pause sites in both *mfd* and Mfd-overexpressing cells, including the *his*, *trp,* and *ops* operators. There were no indications of termination at the pause sites for all locations examined (data not shown). Taken together, our data strongly suggests a model by which Mfd is directly affecting transcription at or near transcription start sites.

5.3 Discussion

In this Chapter we have presented evidence for a role for Mfd outside of DNA repair, showing that Mfd affects transcription in normal wildtype cells. We found that cells overexpressing Mfd have severe growth defects, and overexpressing Mfd at high levels is lethal. Mfd-overexpressing cells have increased levels of tmRNA tagging even without any exogenous inhibitors, implying that Mfd is capable of prematurely terminating transcription. We looked for aberrant transcription-coupled repair activity due to high Mfd levels, but deleting *uvrA*, the first protein to respond in the nucleotide excision repair pathway, did not relieve the growth sensitivity.

We then looked for alterations in genomic transcription in both *mfd* and Mfd-overexpressing cells to further characterize the Mfd termination activity. We subjected

RNA from MG1655 *mfd* and MG1655 pCA24N-Mfd cells to deep sequencing analysis and compared the resulting read coverage profiles to their respective parent cells. We found 94 genes affected 2-fold or more in the *mfd* cells, representing 2% of annotated genes in MG1655. This is less than other global transcription regulators, such as *dksA* (~7%), ppGpp (6%), and H-NS and Fis (~5% each), suggesting that Mfd has more directed targets for regulation (171,173,174). Of the 94 genes, 75 were downregulated and included ncRNA metabolic process genes, nucleobase, nucleoside and nucleotide interconversion genes, DNA replication, and tRNA genes. The remaining 19 genes were upregulated and half were genes involved in carbohydrate catabolic process.

Many of the genes we found affected by deleting *mfd* are involved in growth rate regulation, such as tRNA, RNA modifiers, and DNA replication. Several of the tRNA genes upregulated in *mfd* cells are located in rRNA operons. Unfortunately, the 23S rRNA, 16S rRNA, and 5S rRNA are removed in the library prep of RNAseq due to their high abundance compared to mRNA, so we were unable to see if these RNAs are reduced in *mfd* cells. However, it is possible the entire operon is upregulated in *mfd* cells, and if so Mfd activity would then be downregulating growth rate by directly affecting ribosome levels. This could then explain the severe growth defects of Mfd-overexpressing cells. However, further experiments are needed to fully understand Mfd's role in growth rate regulation. For example, it appears that Mfd is affecting global growth rate and not stringent response, as only 4 genes overlap between the genes

repressed by Mfd and those affected by deleting *dksA* or decreasing ppGpp levels. Deep sequencing analysis of RNA from cells at different stages of growth could address this question.

In contrast with this negative regulator role, however, we found that a significant portion of the downregulated genes in *mfd* cells overlapped with genes affected by deleting *dksA* and reducing ppGpp levels, implying there might be some overlap between Mfd's transcriptional regulation and activity of these regulators. One caveat to this comparison is that our samples were collected during elongation phase (OD_{560} 0.5) while the samples in the *dksA* study were collected at the start of stationary phase (OD_{560} 1.5), when stringent response is expressed. We also compared growth rates for *mfd*, *dksA*, and *mfd dskA* cells and found that the double mutants had a slightly decreased growth rates compared to either single mutant (the ratios for mutant/wildtype growth rates are 0.983, 0.850, and 0.814 for *mfd*, *dksA*, and *mfd dksA* cells, respectively), suggesting there might be some redundancy in activities. Furthermore, DksA has been shown to prevent replication-transcription complexes in a manner independent of its transcription initiation activities (which are responsible for the stringent response), although the exact mechanism of this activity is unknown (81). Perhaps this activity is functionally similar to Mfd's ability to stimulate backtracked elongation complexes, and the overlap in genes affected reflects this mutual activity.

We also looked at the effect of overexpressing Mfd on transcriptional patterns and found that nearly a fourth of the genes in MG1655 are affected, with the vast majority of them being downregulated. We also ruled out that these changes were due to just overexpressing a generic protein by comparing the genes affected to those affected by overexpressing chloramphenicol acetyltransferase. While we could not statistically compare which genes are upregulated or repressed, we were able to look at read coverage profiles to determine if any particular location of the chromosome was targeted. The terminus region, in particular the region between *terA* and *terC*, was severely downgregulated in the Mfd-overexpressing cells. One possibility for this targeting is due to altered transcription of proteins involved in terminus region structure, and more quantitative studies of gene expression should shed light on the possibility of this occurring. Another alternative model is that prophages and other horizontally acquired genes in the terminus region are targeted by Mfd activity. Interestingly, there was one gene, *ydbA*, in the terminus region with that was very highly upregulated (Figure 28). This gene is interrupted by an insertion element, and only the downstream portion of the gene is upgregulated. Genes with an insertion element in other locations in the chromosome (specifically *gatR*, *yhcE*, and *yhiS*) were identified that did not have this profile, meaning this effect is not generic to interrupted genes (data not shown). Transcription terminator Rho is also known to silence prophages and horizontally acquired genes (175,176). However, when we compared the genes affected

by Mfd-overexpressing cells with those that are affected by bicyclomyin treatment (which inhibits Rho activity, from (175)), we found no significant overlap between the two systems. While the Rho study was done using microarray and our Mfd overexpression data is not quantitative, it would appear that Mfd and Rho have different termination targets.

We hypothesized that Mfd was terminating transcription at natural pause sites, and yet the read coverage data contradicted this by showing that termination was occurring at or near promoters. We looked specifically for locations in the chromosome where there were sharp drops in read coverage profiles, yet were unable to find any sharp changes that did not occur at transcription start sites in either *mfd* or Mfd-overexpressing cells. When we narrowed the search to sharp drops within genes, we found two genes in the Mfd overexpressing cell lines with the desired characteristic, confirming that our method of searching was accurate. Even if these are *bona fide* cases of transcription termination within genes, our data clearly supports a model where the majority of the influence on genomic expression by Mfd is occurring near promoters.

Whether Mfd is directly or indirectly affecting gene expression at promoters still remains to be answered. One model is that Mfd is terminating transcription just downstream of the promoter due to a secondary protein. For example, it has been shown in *B. subtilis* that Mfd releases RNAP stalled at a CRE binding site 64 bp downstream of the promoter *dra–nupC–pdp* operon (177). There could perhaps be some

sort of protein or sequence causing RNAP to stall just downstream of the promoter in these genes, causing them to be targeted by Mfd. However, in this case we would expect to see the short transcripts produced due to transcription leading to the blocking protein. It is possible that the resulting RNA is rapidly degraded. While this seems unlikely due to the high sensitivity of RNAseq, RT-PCR or other specific studies of genes known to be downregulated by Mfd could rule out this possibility.

Another model is that Mfd is somehow directly affecting polymerase at promoters. While the exact kinetic and mechanistic details of promoter escape *in vivo* are not fully understood, current evidence shows that RNAP does not simply bind to promoters and rapidly begin elongating. Initiation complexes typically undergo several rounds of abortive initiation (12), with the rate of escape depending on multiple factors (12-14). A more recent study using CHIP-chip analysis showed RNAP is bound to several promoters without any transcriptional activity, which they call "poised promoters" (18). However, one question that remains in this model is how Mfd releases RNAP with sigma present. Mfd requires 25 base pairs of DNA template upstream of the stalled RNAP to act on stalled RNAP, and this site is occupied by σ when it is present (9,147). While the question of the presence of the sigma factor remains, it possible that Mfd is somehow affecting promoter escape.

We conclude that Mfd is capable of affecting gene expressing in wildtype cells, providing a novel role for Mfd as a transcription regulator. While further studies are

necessary to determine the specific targets for Mfd activity, it appears that Mfd is targeting the transcription of genes directly or indirectly near promoters.

5.4 Materials and Methods

Materials

Luria–Bertani broth (LB) contained Bactotryptone (10 g l-1), yeast extract (5 g l-1), and sodium chloride (10 g l-1) and was used for all bacterial growth (with appropriate antibiotics for plasmid selection and the indicated additions); Nitrocellulose membranes are from iBlot Gel Transfer stacks (Invitrogen); tmRNA antibodies were a generous gift from Tania Baker (MIT). RecA and LexA antibodies were purchased from Abcam and Pierce, respectively.

***E. coli* strains**

MG1655 [F- lambda- *ilvG- rfb*-50 *rph*-1] was obtained from the *E. coli* Genetic Stock Center (Yale University). MG1655 *mfd* and MG1655 *uvrA* were made by P1 transduction of *mfd::kan* and *uvrA:kan*, respectively, from the Keio collection with the Kanamycin resistance cassette still remaining.

Plasmids

pMFD19 is pIBI25 derivative that carries the *mfd* gene with its natural promoter and confers ampicilian resistance. The plasmid was created by Selby and Sancar (69) and was provided by Dr. Sancar. pCA24N-Mfd is from the ASKA collection of *E. coli* ORF

without the GFP or His tag (102). Mfd is expressed from the $P_{T5\text{-}lac}$ promoter that can be activated by IPTG.

Western Blots for RecA, LexA, and tmRNA tagging

RecA, LexA, and tmRNA tagging levels were analyzed in MG1655 cells containing pCA24N-Mfd. For the RecA and LexA western blots, cells were pregrown overnight in LB media at 37°C in the presence of chloramphenicol (30 μg/mL), and then diluted to an OD_{560} of 0.1. The cells were then grown to an OD_{560} of 0.5 in the same media, when a 2 mL aliquot of cells was pelleted and frozen. The cells for the tmRNA western blots were grown in the same manner except cells were treated with 100 μM IPTG for the indicated times when they reached OD_{560} 0.5. At each timepoint, 2 mL worth of OD_{560} 0.5 cells were removed and frozen. 2 mL of untreated cells were removed at OD_{560} 0.5 as a control.

For all western blots, the frozen cell pellets were thawed at room temperature, resuspended in 25 μL water and 25 μL SDS loading buffer (20% glycerol, 100 mM Tris pH 6.8, 2% SDS, 2% β-mercaptoethanol, and 0.1 mg Bromophenol blue), vortexed, and boiled for 5 minutes. Lysates were loaded onto a 4-20% Mini-PROTEAN TGX gels (Bio-Rad Laboratories) and the gels were run at constant voltage (200 V) for 45 min in 25 mM Tris-Glycine buffer with 0.1% SDS. For the RecA and LexA western blots, protein concentrations in the cell lysates were measured using a BCA Protein Assay Kit (Thermo Scientific) using BSA standards. Gels were transferred to a nitrocellulose membrane for

60 min at 12 V using a Genie Blotter transfer device (Idea Scientific Co.). The blot was blocked for 1 h in 20% non-fat milk powder solution (Biorad) in Tris-buffered saline (TBS). The membrane was incubated overnight at 4°C with either polyclonal RecA, LexA, or tmRNA tag primary antibody and Tween (0.1%), and then washed three times with TBS buffer at room temperature (10 min each). The membrane was incubated with secondary antibody (IRDye 800CW-conjugated goat anti-rabbit IgG (LI-COR®)) for 30 min, and the washes were repeated. After air-drying, the membrane was scanned on an Odyssey Imaging System (LI-COR Biosciences), and quantified using the Odyssey software (version 3.0).

Growth kinetics for *uvrA* pMfD19 cells

Overnight cultures of MG1655 and MG1655 *uvrA* cells containing pMfd19 were pregrown in LB containing ampicillin (50 μg/mL). Cells were diluted to an OD_{630} of 0.5, diluted 1:200 in LB with ampicillin, and then mixed with an equal volume (75 μl each) of LB containing serial dilutions of Stl (starting at 4.5 μg/ml) in 96-well plates. The cells were grown in a BioTek ELx808 Microplate Reader at 37°C with constant shaking, and OD_{630} was measured every 15 minutes for 18 h. Cell growth was analyzed as in previous chapters.

Deep sequencing analysis

Three independent isolates of MG1655 and MG1655 *mfd* cells were grown to approximately mid-log phase (approximately OD_{560} 0.5) in LB, when 2 mL worth of

OD_{560} 0.5 cells were collected and frozen. For the Mfd overexpressing experiments, MG1655 pCA24N and MG1655 pCA24N-Mfd cells were grown to roughly OD_{560} 0.25 and treated with 100 mM IPTG for 1 hour. Cell pellets were lysed and RNA collected using Qiagen's RNeasy Plus Mini Kit with Qiagen Bacteria Protect RNA kit. RNA samples were then treated with DNase (New England Biolabs) for 30 min at 37 °C. Reactions were stopped by adding 5 mM EDTA and incubating at 75 °C for 10 minutes.

Library preparation, rRNA depletion, and deep sequencing were performed at the Sequencing Facility of the Duke Institute for Genome Sciences & Policy (IGSP). rRNA was depleted using Ribo-Zero™ rRNA Removal Kits (Epicentre). Deep sequencing was performed via Illumina HiSeq™ (50 bp single reads). The sequencing reads (approximately 13 to 25 million per sample) were imported into Geneious Pro (Biomatters) and assembled to the reference chromosome MG1655 (GenBank Accession Number 000913.2). The assembly process was set to medium/low sensitivity on Geneious, with the following parameters: 10% gaps allowed per read; word length of 18; index word length of 13; words repeated more than 12 times ignored; 20% maximum mismatches per read; and maximum ambiguity of 4. Read coverage maps and RPKM data was subsequently generated by Geneious. BAM files of the resulting assembly data were exported to JMP Genomics (SAS). TMM normalization and ANOVA analysis of the read samples were conducted in JMP.

Sharp changes in read coverage

To locate sharp read changes in the Mfd-overexpressing cells, we first eliminated genes that were less than 400 bp long (too small for our analysis) and had RPKM values less than 2.5 (poor expression). We then took the transcription start and end values and rounded to the nearest hundred basepair to allow for alignment with the 100 bp bin RPKM values. We also added/subtracted 100 bp from the ends to minimize expression overlaps between genes. We then took the ratio of the beginning 100 bp RPKM value to the end value for each gene in each cell line, accounting for gene directionality. The resulting ratios for the Mfd-overexpressing cell lines were then divided by the WT ratios to identify genes with sharp drops in expression levels only in the Mfd-overexpressing cell lines.

6. Conclusions and Future Directions

6.1 Summary of Results

In this dissertation, we examined the fate of transcription elongation complexes stalled by DNA-protein complexes, Streptolydigin, and Actinomycin D. Transcription is a vital process for protein production and cell survival, and therefore it is important to know how the cell responds to these types of stress. Stalled transcription complexes can also lead to replication-transcription collisions, which can lead to DNA damage, genomic rearrangements, and cell death (31,70,81). The cell must therefore have a way to clear these complexes, either by restarting them or removing them from DNA. Understanding how the cell responds to blockages in transcription provides insight into the requirements for productive elongation. Furthermore, several antibiotics function by blocking transcription, and understanding how the cell responds to these inhibitors could give insights on how to make the inhibitors more effective.

DNA-protein crosslinks form when proteins become covalently trapped on DNA. The type of proteins and the locations where they form all vary, making DPC study quite difficult. As a result, they are one of the more poorly understood forms of DNA damage. We use 5-azacytidine as a model system to combat the inherent variability of DPC study. Aza-C is a cytidine analog that covalently traps cytosine methyltransferases at known recognition sequences. Using this system, we showed that the tmRNA system is required for cell survival during DPC formation due to its ability

to recycle stalled ribosomes. This raised the question of how tmRNA is able to access the stalled ribosome. Previous members of our lab looked at tmRNA tagging levels of aza-C hypersensitive mutants, but did not identify and mutants with decreased tagging that would be expected to release RNAP. In this study, we looked for evidence of A-site cleavage by testing candidate *rnase* knockout mutants, but again did not find a role for A-site cleavage in cell survival during DPC formation. The mechanism for tmRNA activity during DPC formation is still unknown.

The aza-C hypersensitivity screen conducted previously identified *uvrD* and *dinG* knockouts as hypersensitive to aza-C, potentially indicating that these are the factors responsible for clearing stalled RNAP. We also discovered that *dksA* knock-out mutants are slightly resistant to Aza-C treatment, and have increased levels of tmRNA tagging. DksA is an elongation factor that could be competing with the release mechanism responsible for tmRNA activity. These results expanded upon our previously proposed "chain-reaction" model (See Figure 34).

Using the same genetic screening and candidate mutant approach, we also attempted to identify the mechanism for the repair of DPCs, and ruled out many pathways that could be expected to be involved. We did find that *dnaK* mutants are resistant to aza-C treatment. Taken together with the previous discovery that *dnaJ* mutants are hypersensitive, we propose that there could be some DnaJ-dependent, DnaK-independent protein degradation involved in DNA repair.

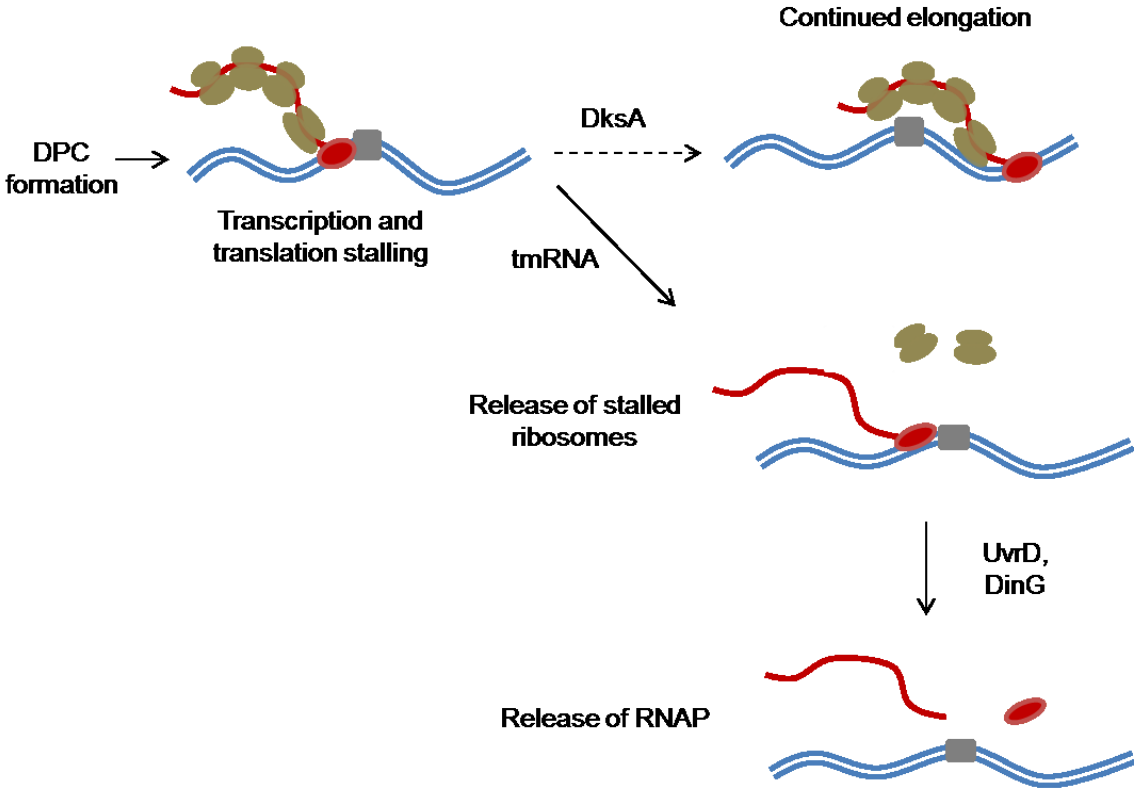


Figure 34: Model for the fate of elongation complexes at DPCs
DPC formation leads to transcriptional and translational stalling. The stalled RNAP can either be cleared most likely through the activities of UvrD and DinG, or it can potentially continue elongation with the assistance of DksA. The tmRNA system is also responsible for clearing the stalled ribosomes in an A-site cleavage-independent manner. The model shown here has tmRNA clearing the ribosomes before the release of RNAP, but the exact order of these steps has not been elucidated.

Because the study of the effect of DPCs on transcription is complicated by the other uncertainties inherent in DPC formation (such as how DPCs are repaired), we used the elongation inhibitors Streptolydigin and Actinomycin D to isolate the effects on transcription. We showed that Stl binding is transient and suggest that RNAP needs to be held in an open conformation for efficient inhibition. This latter finding was supported by the discovery that a mutant polymerase with increased pausing is hypersensitive to Stl, while a mutant polymerase with decreased pausing is resistant. We also found that a mutant polymerase with decreased pausing is dramatically resistant to ActD, providing surprising evidence for the role of RNAP dynamics in ActD sensitivity. We looked for additional Rif-resistant mutations in the β subunit of RNAP that lead to ActD resistance, and identified 8 mutations in 6 positions that lead to resistance. This data suggests a novel mechanism for ActD inhibition of RNAP that is more complicated than just steric hindrance on DNA.

With our increased understanding of Stl and ActD inhibition, we were able to address the question of the fate of transcription complexes stalled by exogenous inhibitors. Contrary to DPCs, the transcription-coupled repair pathway appears to clear stalled elongation complexes, competing with productive elongation that could occur once the drug binding reverses. We also provided evidence indicating that Mfd removal of elongation complexes creates the substrates necessary for the involvement of the tmRNA system. Mfd termination was also shown to be responsible for part of the

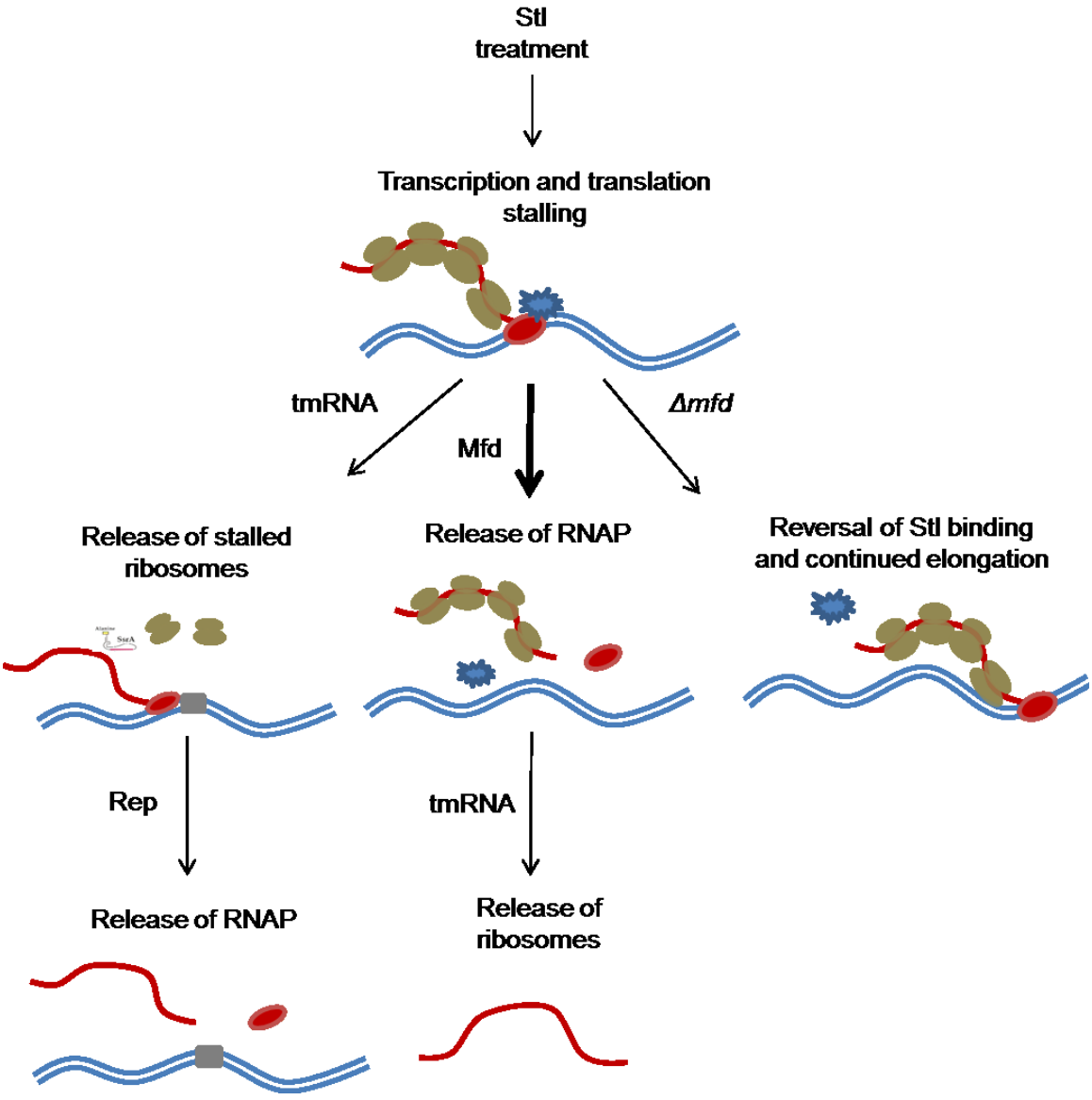


Figure 35: Model for the fate of elongation complexes stalled by Stl
During Stl treatment, Mfd is responsible for releasing the majority of Stl-stalled transcripts. This creates substrates for the tmRNA system, which then recycles the stalled ribosomes on the nascent transcript. The tmRNA system is also capable of releasing a small amount of ribosomes without Mfd activity. Rep also releases a small amount of RNAP, presumably ahead of replication forks. In the absence of Mfd, the transient nature of Stl binding allows for the proper resumption of the majority of the elongation complexes.

hypersensitivity of *rpoB8* mutants, potentially due to the increased amount of paused polymerases that are targets for Mfd termination. Taken together, we propose that Mfd termination and productive elongation are in competition with one another during elongation inhibitor treatment. See Figure 35 for our proposed model.

We also propose a novel role for Mfd in regulating global gene expression when there is no DNA damage or exogenous inhibitors present. Cells overexpressing Mfd were shown to have increased levels of tmRNA tagging, which implies Mfd activity is leading to transcription termination either directly or indirectly. This termination does not appear to be due to aberrant NER repair. If Mfd is terminating elongation, it could be expected to be altering global transcription patterns, and therefore we subjected RNA isolated from *mfd* and Mfd-overexpressing cells to RNAseq. In the *mfd* knockout cells we found that genes involved in ncRNA metabolic processes, nucleotide interconversion, and DNA replication were upregulated. Many tRNA genes, in particular several located in rRNA operons, were also upregulated. Genes involved in catabolic processes were downregulated, several of which overlapped with genes that are downregulated in *dksA* and ppGpp° cells. Taken together, the RNAseq results implied that Mfd might play a role in growth rate regulation.

In the Mfd-overexpressing cells, nearly a fourth of the annotated genes in MG1655 were affected, with the vast majority of these affected genes downregulated. The Terminus region was particularly underrepresented in Mfd-overexpressing cells.

This data supported our model that Mfd activity results in transcription termination. With the discovery that Mfd negatively affects cells expressing an RNAP mutant with increased pausing, we hypothesized that Mfd is terminating transcription at natural pause sites. Surprisingly, analysis of the location of deviation between *mfd* and wildtype read coverage showed that transcription profiles deviate at or near promoters at nearly all locations examined. We propose instead that Mfd is either directly or indirectly affecting promoter escape.

6.2 "Chain-reaction" model for DPC consequences

In Chapter 2 we introduced the "chain-reaction model" for the effects on DPC formation on translocating proteins. Our lab previously showed the replication forks are inhibited by DPCs (42), and here we show genetic evidence that DPCs block transcription and translation. The tmRNA system is recruited to remove the stalled polymerase, while the RNAP is cleared via an unknown mechanism. Presumably, once the stalled translation and elongation complexes are cleared the DPC itself can be repaired.

The most likely candidates for RNAP release during DPC formation are DinG and UvrD. Transposon insertions in these genes resulted in hypersensitivity to aza-C, and these proteins have been shown to promote cell survival during replication/transcription collisions (70). Furthermore, UvrD has been implicated as

having a general role in removing proteins from DNA to assist with the translocation of replication forks (71). Conversely, DksA activity appears to be detrimental to cell survival during DPC formation, which could potentially be related to its competition with tmRNA recycling. Further genetic tests with DinG, UvrD, DksA, and SmpB in various double-mutant combinations could shed light on the interplay of these pathways. *In vitro* elongation studies would also be useful in confirming the details of this model. First, assays utilizing a template with a DPC (similar to the one used in (43)) could be employed to look for proteins capable of releasing RNAP, including DinG and UvrD. The use of DksA in these studies could also show if it is capable of promoting transcription past the lesion, or if it competes with release if a release protein is confirmed.

In this model, we have yet to discern whether the release of the ribosome or the RNAP occurs first. If RNAP is released first, the nascent transcript would also be released, allowing translation to continue to the end of the transcript and tmRNA to bind to the empty A-site. However, tmRNA has been shown to rescue ribosomes at rare codons, indicating that there is a mechanism for tmRNA activity when ribosomes are stalled in the middle of a transcript (88). Given that ribosome pausing stimulates both A-site cleavage and tmRNA activity, and that the first step of tmRNA rescue is binding to the empty A-site of a ribosome, it was proposed that these two pathways were coupled (90). However, it has since been shown that RNase II activity is required for

A-site cleavage (124), yet a recent study by the same lab showed that tmRNA tagging levels are unaffected in RNase II deficient cells (90). Therefore, a mechanism must exist in the cells for tmRNA to act on ribosomes stalled in the middle of a transcript that is independent of RNase II–mediated A-site cleavage.

6.3 Repair of DPCs

The repair of DPCs has been an ongoing topic of interest in our lab. While it was not a direct aim of this dissertation, genetic tests were conducted in an effort to at minimum identify potential repair pathways. To this end, we found that *dnaK* mutants were resistant to DPC formation. When taken together with the discovery that *dnaJ* mutants are hypersensitive, these data present two possible pathways for the role of these chaperone proteins: DnaJ might be repairing DPCs directly, with DnaK partially competing with this activity, or the induction of heat shock proteins in *dnaK* mutants is facilitating repair and DnaJ assists in cell survival in another manner. Further experiments are needed to separate these two possibilities.

To identify if DnaJ is affecting DPC repair directly, an assay measuring the level of DPCs *in vivo* would be of great use. A previous member of our lab presented a promising Western-based assay for monitoring DPC formation *in vivo,* showing that he could detect DNA-bound MTase in a western blot, running slower than free MTase (Kenny Kuo, unpublished results). Once fully developed, this assay could be used to

monitor the level of DPC formation in *dnaJ* knockout cells. Additionally, the 2D-gel based assay used to show that DPCs block replication (42) could be used to monitor the blockage of replication in *dnaJ* and *dnaK* cells, as alterations in the levels of blocked replication forks would indicate changes in the efficiency of DPC repair.

To test if stimulation of heat shock proteins is the cause for the resistance to aza-C in *dnaK* mutants, the effects of DPC formation during a 42 °C pulse treatment (to induce heat shock expression) could be tested. Conversely, the heat shock response can be initiated by overexpressing RpoH (sigma 32) (178), so aza-C hypersensitivity profiles for cells overexpressing RpoH could also answer if the heat shock response is responsible for cell survival during DPC formation.

6.4 Actinomycin D mechanism of inhibition

To understand the mechanism of ActD inhibition, we would like to characterize mutations that lead to ActD resistance. However, we found a surprisingly high rate of mutation frequency leading to ActD resistance, implying there might be a mechanism for overcoming ActD entry into the cell even when the membrane is compromised. This system could potentially play a role in the mechanisms of ActD resistance, or it could be a more trivial way for the cell to not be affected by the drug. One possibility in the latter scenario is that the drug is being pumped out via an efflux system. Therefore we would like to test knockouts of different efflux systems for their sensitivity to ActD. A recent

study identified 20 different proteins belonging to five families that are capable of transporting various classes of antibiotics and small molecules (179). We would like to test knockouts of these genes to see if we can identify the system(s) responsible for the ActD resistance.

The intercalation of Actinomycin D into DNA has been well established (63,64). However, the fact we were able to isolate *rpoB* mutants that are resistant to ActD suggests that there could potentially be direct interactions between ActD and RNAP once RNAP transcribes to the site of ActD intercalation. We would be greatly interested in seeing if ActD binds directly to RNAP, or if it is the properties of the RNAP mutants that lead to resistance to ActD. One potential assay for this would be to test various RNAP mutants, in particular RpoB2 and RpoB8, in an *in vitro* system using ActD treatment. We could monitor the transcription rates for each mutant in the absence of ActD, then see how transcription varies when ActD is added. If RNAP and ActD are interacting directly, you would expect there would be multiple possible substitutions in the putative ActD binding site of RNAP that could lead to ActD resistance. Additionally, more comprehensive screening of *rpoB* mutants could further detail the putative ActD binding site.

6.5 Factors that recognize Stl-stalled RNAP

Unlike RNAP stalled by DPCs, Stl-stalled RNAP is primarily cleared by the the transcription-coupled repair pathway. However, our genetic studies also indicated that Rep activity also inhibits cell survival during Stl treatment. Like DinG and UvrD, Rep is capable of releasing RNAP when replication forks collide (70,71). Rep travels with the replication fork and acts as a "bulldozer" to remove proteins that would otherwise impede translocation (71). We propose that Mfd is the primary factor that removes Stl-stalled RNAP, while Rep removes a small fraction when replication forks collide with the immobile RNAP. Similar to DinG and UvrD, this activity does not appear to lead to tmRNA activity, and instead another mechanism must be employed for tmRNA recycling.

We attempted to analyze drug sensitivity profiles with cells overexpressing Rep, but found that full expression of Rep from the pCA24N plasmid with IPTG induction was lethal. A cleaner expression system is needed to be certain Rep activity is detrimental to cell growth during Stl treatment. One option is the pBad plasmid system, which is fully titratable in specially designed cell lines such as BW27783 (98,180). *In vitro* studies looking for Rep termination could also be conducted in a similar manner to Mfd (see below) to see if Rep is capable of releasing Stl-stalled RNAP, although it might need to be part of a replication fork to be active.

6.6 In vitro elongation studies with Streptolydigin and Actinomycin D

In vitro elongation studies can help give insight into both the mechanism of inhibition and the factors capable of releasing stalled RNAP. As mentioned previously, they can be applied to all three methods of inhibition used in this dissertation. However, certain technical challenges were presented in the previous chapters that need to be addressed for these types of assays.

For reactions involving Stl, we must first identify the conditions in which Stl is capable of inhibiting RNAP. Using the PCR fragment from pDE13 as a template, Erie *et al.* showed that nucleotide starved RNAP undergoes conformation switches between the activated and inactivated states as it attempts to undergo further rounds of nucleotide addition (136). While addition of a proper nucleotide proceeds rapidly, the misincorporation of an incorrect nucleotide occurs at a much slower rate, which could potentially give Stl enough time to bind to an inactivated state ((136) and Dorthy Erie, personal communication). Therefore a preincubation step with Stl could lead to inhibition of elongation in our current assay. To expand upon our hypothesis that Stl binding requires a paused or otherwise arrested RNAP, we could use a template with a well-studied pause site and see if Stl is capable of stalling RNAP on this template.

In order to study the role of transcription factors such as Mfd in removing stalled RNAP, we propose expanding the current studies to using a bead-based assay to directly monitor transcript release. These assays are based on the work of Park *et al.*

(147), who used DNA templates bound to a magnetic bead to separate RNA transcript products into "pellet" (still in the ternary complex and thus attached to the magnetic beads) and "supernatant" (released from the complex) fractions. In this manner, we could study the release of transcripts under conditions that do not promote full-length transcript production, such as in the case of a tight ActD complex.

The proposed method is as follows: PCR of the pDE13 fragment is conducted using a biotin labeled primer, and the resulting products are pre-bound to streptavidin-coated magnetic beads. +24 complexes are formed in a similar manner as previously, including a pre-incubation step with ActD or Stl. All four nucleotides are then added to allow for elongation to continue if possible, assuring that complex stalling is now to due the presence of the inhibitor and not nucleotide starvation. A long incubation with Mfd (or another potential release factor) could then be implemented, and the reactions separated into pellet and supernatant fractions by magnetically capturing the beads. We could then monitor the presence of transcripts smaller than the full-length product in the pellet or supernatant fractions, with the presence of transcripts in the supernatant indicating release. This assay would allow for direct monitoring of release, instead of simply inferring release from the lack of full-length transcripts.

6.7 Identifying the targets of Mfd termination in wildtype cells

We were surprised to find that cells overexpressing Mfd had high levels of tmRNA tagging even in the absence of Stl. This implies that Mfd has termination activity in wildtype cells, which is further supported by the growth deficiencies of Mfd-overexpressing cells. However, our expression system is one that expresses Mfd at very high levels in an on-or off- manner. More careful expression of Mfd would help separate out potential secondary effects of recombinant protein expression, such as by cloning Mfd into a pBad plasmid and expressing in BW27783 (see above).

While not quantitative, the RNAseq data with the Mfd-overexpressing cells showed that cells undergo vast transcriptional modifications when Mfd activity is elevated. Higher sample numbers would allow for statistical analysis of the genes affected, and could be useful in parsing out specific targets for Mfd activity. In addition, we would like to introduce a control that isolates the effect of Mfd activity on gene expression. Chambers *et al.* identified a point mutation in the TRG (translocase in RecG) domain that eliminates Mfd's ability to release RNAP but not its ATPase or DNA binding abilities (125). RNAseq studies using this mutant would serve as a solid control to remove genes affected by protein overexpression.

The RNAseq results from the *mfd* mutant cell lines indicated a possible role for Mfd in global growth rate regulation. Of particular interest are the changes to rRNA and tRNA levels, as these are the primary genes that regulate growth rate. However, we

were unable to monitor changes in rRNA levels using RNAseq as these RNAs are removed during library formation. RT-PCR of these genes from *mfd* cells could confirm the effect of Mfd on tRNA genes located in ribosomal operons, and further indicate if Mfd is affecting levels of rRNAs. RNAseq analysis of *mfd* cells under different growth conditions, such as at different temperatures and different stages of growth could also help address if Mfd is affecting global growth rate.

We also proposed that Mfd is affecting gene expression at promoters, potentially by affecting promoter escapes. It is tempting to speculate that Mfd targets RNAP discovered to be "poised" at promoters (18). To test this, ChIP-Seq experiments monitoring the location of RNAP in *mfd* mutants compared to wildtype cells could indicate if Mfd is affecting the distribution of RNAP in cells.

6.8 Final Remarks

The initial goal of this dissertation was to identify the fate of transcription elongation stalled at DNA damage in the form of DNA-protein crosslinks. While we have identified candidate factors involved in recognizing stalled RNAP, the study of the effects of DPCs on elongation is complicated by the numerous primary and secondary effects of DPC formation; replication has been shown to be blocked by DPCs (42), and could also be colliding with RNAP that is already trapped at a DPC. Repair factors attempting to repair the DPC might also be present, resulting in a potentially large "pile-

up" at DPCs. We had hoped that the use of Stl and ActD would provide a way to isolate the effects on transcription, particularly if the same pathway responded in both systems. However, our results imply that clearing of stalled RNAP complexes is more complex, and that different pathways can respond under different conditions.

During DPC formation, the tmRNA system is responsible for clearing the stalled ribosomes in an A-site cleavage-independent manner. The RNAP is also cleared most likely through the activities of UvrD and DinG. DksA also assists in cell survival in a way that is counteractive to the tmRNA system. This suggests that perhaps some transcription can occur past the lesion, though this is still highly speculative. Notably, the transcription-coupled repair pathway does not assist in cell survival during DPC formation. Given that Mfd is capable of releasing RNAP stalled at a bound protein in many other conditions (66,125,177), it appears that the activities of UvrD and DinG outcompete Mfd for RNAP binding.

In the case of Stl and ActD inhibition, we propose that Mfd is responsible for releasing stalled polymerase, but this release causes a premature termination phenotype and is detrimental to cell growth. This activity creates substrates for the tmRNA system, which accounts for part of the total tmRNA activity in the cell. tmRNA is apparently capable of releasing a small fraction of ribosomes in the absence of Mfd, perhaps in a similar manner as with DPC formation. Rep also removes Stl-stalled RNAP, presumably ahead of a colliding replication fork. This most likely occurs after tmRNA

has already cleared the ribosomes, as Rep does not create substrates for the tmRNA system. In the absence of Mfd, the transient nature of Stl binding allows for proper elongation to continue for the majority of elongation complexes. tmRNA is apparently capable of releasing ribosomes during the short amount of time that they are stalled, as tmRNA contributes to cell survival in the absence of Mfd.

This dissertation confirms previous studies that show the importance of transcription regulation in cell survival during stressful situations. The cell has numerous, partially redundant pathways to help clear stalled transcription complexes and maintain a pool of productive RNAP via recycling. Certain antibiotics can take advantage of these pathways, turning this recycling of RNAP into detrimental premature termination. Elucidating the cellular consequences of transcriptional arrest assists in both the understanding of an essential cellular process and increases the potential for better antibiotic design and application.

Appendix I

Gene	Function
dinG	helicase, prevents replication/transcription collisions
dksA	transcription elongation factor; assists with stringent response
dnaJ	Chaperone
dnaK	Chaperone
greA	elongation factor; stimulates mRNA cleavage activity of RNAP
greB	elongation factor; stimulates mRNA cleavage activity of RNAP
hepA	RNAP recycling factor
hflC	component of hflCK complex; regulates FtsH protease
lexA	transcription repressor; suppresses the SOS response
mfd	transcription-coupled repair factor
Mu 9	unknown function
Mu Gam	nuclease inhibitor
nusA	transcription terminator/antiterminator factor
recA	functions in DNA strand exchange and recombination
recB	helicase; essential for recombination and dsDNA break repair
recC	essential for recombination and dsDNA break repair
recF	functions in RecA-mediated recombination
recG	helicase; catalyzes branch migration on forked DNA structures
rep	helicase; removes proteins from DNA ahead of replication fork
rho	transcription termination factor
rnase D	exonuclease involved in the 3' ribonucleolytic processing of precursor tRNA
rnaseII	3' exonuclease; involved in tRNA processing
rnaseLS	endoribonuclease that plays a role in mRNA degradation
rnasePH	3' exonuclease; involved in tRNA processing
sbcCD	exonuclease complex; inserts double-strand breaks in protein-bound DNA
smpB	tmRNA cofactor; required for stable tmRNA binding to ribosomes
ssrA	specialized tmRNA; ribosome recycling
trmH	methyltransferase; methylates a subset of tRNAs and rRNAs
umuCD	translesion polymerase involved in SOS response
uvrA	subunit of the UvrABC nucleotide excision repair complex
uvrD	helicase; prevents replication/transcription collisions
yejH	putative helicase; complements UV repair damage

References

1. Landick, R., and Yanofsky, C. . (1987) In F. C. Neidhardt, J. L. I., K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coil and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington D.C., pp. 1276-1301.

2. Landick, R., and Turnbough, C. L. (1992) In Yamamoto, S. L. M. a. K. R. (ed.), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp. 407-446.

3. Roberts, J.W., Shankar, S., Filter, J.J. (2008) RNA polymerase elongation factors. *Annu. Rev. Microbiol.*, **62**, 211-233.

4. Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K. and Darst, S.A. (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A° resolution. *Cell*, **98**, 811–824.

5. Cramer, P., Bushnell, D.A. and Kornberg, R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*, **292**, 1863-1876.

6. Minakhin, L., Bhagat, S., Brunning, A., Campbell, E.A., Darst, S.A., Ebright, R.H. and Severinov, K. (2001) Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc Natl Acad Sci U S A*, **98**, 892–897.

7. Werner, F. (2008) Structural evolution of multisubunit RNA polymerases. *Trends in Microbiology*, **16**, 247-250.

8. Mooney, R.A., Darst, S.A. and Landick, R. (2005) Sigma and RNA polymerase: An on-again, off-again relationship? *Mol Cell*, **20**, 335–345.

9. Marr, M.T., Datwyler, S.A., Meares, C.F. and Roberts, J.W. (2001) Restructuring of an RNA polymerase holoenzyme elongation complex by lambdoid phage Q protein. *Proc Natl Acad Sci U S A*, **98**, 8972–8978.

10. Ring, B.Z., Yarnell, W.S. and Roberts, J.W. (1996) Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*, **86**, 485-493.

11. Mooney, R.A. and Landick, R. (2003) Tethering sigma70 to RNA polymerase reveals high in vivo activity of sigma factors and sigma70-dependent pausing at promoter-distal locations. *Genes Dev*, **17**, 2839-2851.

12. Goldman, S.R., Ebright, R.H. and Nickels, B.E. (2009) Direct Detection of Abortive RNA Transcripts in Vivo. *Science*, **324**, 927–928.

13. Stepanova, E., Lee, J., Ozerova, M., Semenova, E., Datsenko, K., Wanner, B.L., Severinov, K. and Borukhov, S. (2007) Analysis of promoter targets for *Escherichia coli* transcription elongation factor GreA In vivo and in vitro. *J Bacteriol*, **189**, 8772–8785.

14. Gourse, R.L., Gaal, T., Bartlett, M.S., Appleman, J.A. and Ross, W. (1996) rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu Rev Microbiol*, **50**, 645-677.

15. Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H. and Artsimovitch, I. (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*, **448**, 157-162.

16. Bar-Nahum, G. and Nudler, E. (2001) Isolation and characterization of sigma(70)-retaining transcription elongation complexes from Escherichia coli. *Cell*, **106**, 443-451.

17. Mukhopadhyay, J., Kapanidis, A.N., Mekler, V., Kortkhonjia, E., Ebright, Y.W. and Ebright, R.H. (2001) Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell*, **106**, 453–463.

18. Reppas, N.B., Wade, J.T., Church, G.M. and Struhl, K. (2006) The transition between transcriptional initiation and elongation in *E. coli* is highly variable and often rate limiting. *Mol. Cell*, **24**, 747–757.

19. Vogel, U. and Jensen, K. (1995) The RNA chain elongation rate in Escherichia coil depends on the growth rate. *J. Bacteriol.*, **176**, 2807-2813.

20. Liu, K., Zhang, Y., Severinov, K., Das, A. and Hanna, M.M. (1996) Role of Escherichia coli RNA polymerase alpha subunit in modulation of pausing, termination and anti-termination by the transcription elongation factor NusA. *EMBO J.*, **15**, 150–161.

21. Nudler, E. (1999) Transcription elongation: structural basis and mechanisms. *J. Mol. Biol.*, **288**, 1–12.

22. Erie, D.A. and Kennedy, S.R. (2009) Forks, pincers, and triggers: the tools for nucleotide incorporation and translocation in multi-subunit RNA polymerases. *Current Opinion in Structural Biology*, **19**, 708–714.

23. Toulokhonov, I., Zhang, J., Palangat, M. and Landick, R. (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell*, **27**, 406–419.

24. Erie, D.A. (2002) The many conformational states of RNA polymerase elongation complexes and their roles in the regulation of transcription. *Biochimica et Biophysica Acta*, **1577**, 224– 239.

25. Erie, D.A., Yager, T.D. and Hippel, P.H.v. (1992) The single-nucleotide addition cycle in transcription: a Biophysical and Biochemical perspective. *Annu. Rev. Biophys. Biomol. Struct.*, **21**, 379-415.

26. Temiakov, D., Patlan, V., Anikin, M., McAllister, W.T., Yokoyama, S. and Vassylyev, D.G. (2004) Structural basis for substrate selection by T7 RNA polymerase. *Cell*, **116**, 381–391.

27. Landick, R. (2006) The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.*, **34**, 1062–1066.

28. Artsimovitch, I. and Landick, R. (2000) Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 7090–7095.

29. Toulokhonov, I., Artsimovitch, I. and Landick, R. (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science*, **292**, 730–733.

30. Borukhov, S., Lee, J., Laptenko, O. (2005) Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol. Microbiol.*, **55**, 1315-1324.

31. Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M.E., and Nudler, E. (2011) Linking RNA Polymerase Backtracking to Genome Instability in E. coli. *Cell*, **146**, 533–543.

32. Kireeva, M.L. and Kashlev, M. (2009) Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 8900–8905.

33. Neuman, K.C., Abbondanzieri, E.A., Landick, R., Gelles, J. and Block, S.M. (2003) Ubiquitous Transcriptional Pausing Is Independent of RNA Polymerase Backtracking. *Cell*, **115**, 437–447.

34. Lesnik, E.A., Sampath, R., Levene, H.B., Henderson, T.J., McNeil, J.A. and Ecker, D.J. (2001) Prediction of rho-independent transcriptional terminators in *Escherichia coli. Nucleic Acids Res.*, **29**, 3583–3594.

35. Wilson, K.S. and Hippel, P.H.v. (1995) Transcription termination at intrinsic terminators: The role of the RNA hairpin. *Proc. Natl. Acad. Sci. U. S. A.*, **92**, 8793-8797.

36. Nudler, E. and Gottesman, M.E. (2002) Transcription termination and anti-termination in *E. coli*. *Genes to Cells*, **7**, 755–768.

37. Peters, J.M., Vangeloff, A.D. and Landick, R. (2011) Bacterial Transcription Terminators: The RNA 3′-End Chronicles. *J. Mol. Biol.*, **412**, 793–813.

38. Pomerantz, R.T., and O'Donnell, M. (2008) The replisome uses mRNA as a primer after colliding with RNA polymerase. *Nature*, **456**, 762–766.

39. Pomerantz, R.T., and O'Donnell, M. (2010) Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science*, **327** 590–592.

40. Nakano, T., Morishita, S., Katafuchi, A., Matsubara, M., Horikawa, Y., Terato, H., Salem, A.M.H., Izumi, S., Pack, S.P, Makino, K., Ide, H. (2007) Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks. *Mol. Cell*, **28**, 147-158.

41. Barker, S., Weinfeld, M., Murrary, D. (2005) DNA-protein crosslink: their induction, repair, and biological consequences. *Mutational Research*, **589**, 111-135.

42. Kuo, H.K., Griffith, J.D. and Kreuzer, K.N. (2007) 5-Azacytidine induced methyltransferase-DNA adducts block DNA replication *in vivo. Cancer Res*, **67**, 8248-8254.

43. Som, S. and Friedman, S. (1994) Inhibition of transcription *in vitro* by binding of DNA (cytosine-5)-methylases to DNA templates containing cytosine analogs. *J. Biol. Chem.*, **269**, 25986-25991.

44. Christman, J.K. (2002) 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene*, **21**, 5483-5495.

45. Santi, D.V., Norment, A., Garrett, C.E. . (1984) Covalent bond formation, between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. *Proc. NatI. Acad. Sci.*, **81**, 6993-6997.

46. Som, S., Friedman, S. (1994) Inhibition of Transcription *in vitro* by Binding of DNA(Cytosine-5)-Methylases to DNA Templates Containing Cytosine Analogs. *Journal of Biological Chemistry*, **269**, 25986-25991.

47. Bhagwat, A.S., Roberts, R.J. (1987) Genetic Analysis of the 5-Azacytidine Sensitivity of *Escherichia coli* K-12. *Journal of Bacteriology*, **169**, 1537-1546.

48. Klimasauskas, S., Kumar, S., Roberts, R.J., Cheng, X. (1994) Hhal Methyltransferase Flips Its Target Base Out of the DNA Helix. *Cell*, **76**, 357-369.

49. Klimasauskas, S., Szyperski, T., Serva, S., Wuthrich, K. (1998) Dynamic modes of the flipped-out cytosine during HhaI methyltransferase–DNA interactions in solution. *EMBO Journal*, **17** 317–324.

50. O'Gara, M., Klimasauskas, S., Roberts, R.J., Cheng, X. (1996) Enzymatic C5-Cytosine Methylation of DNA: Mechanistic Implications of New Crystal Structures for HhaI Methyltransferase-DNA-AdoHcy Complexes. *J. Mol. Biol.*, **261**, 634–645.

51. Gabbara, S., Bhagwat, A.S. (1995) The mechanism of inhibition of DNA (cytosine-5-)-methyltransferases by 5-azacytosine is likely to involve methyl transfer to the inhibitor. *Biochem. J.*, **307**, 87-92.

52. Friedman, S. (1986) Binding of the EcoRII methylaes to azacytosine-containing DNA. *Nucleic Acids Research*, **14**, 4543-4556.

53. Salem, A.M.H., Nakano, T., Takuwa, M., Matoba, N., Tsuboi, T., Terato, H., Yamamoto, K., Yamada, M., Nohmi, T., Ide, H. (2009) Genetic analysis of repair and damage tolerance mechanisms for DNA-protein cross-links in *Escherichia coli*. *J. Bacteriol.*, **191**, 5657-5668.

54. Garcia, J.S., Jain, N., Godley, L.A. (2010) An update on the safety and efficacy of decitabine in the treatment of myelodysplastic syndromes *Onco. Targets and Therapy*, **3**.

55. Glover, A.B., Leyland-Jones, B.R., Chun, H.G., Davies, B., and Hoth, D.F. . (1987) Azacitidine: 10 years later. *Cancer Treatment Reports*, **71**, 737-746.

56. Baylin, S.B., Esteller, M., Rountree, M.R., Bachman, K.E., Schuebel, K., Herman, J.G. (2001) Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Human Molecular Genetics*, **7**, 687-692.

57. Siddhikol, C., Erbstoeszer, J.W., and Weisblum, B. (1969) Mode of Action of Streptolydigin. *J. Bacteriol.*, **99**, 151-155.

58. Von Meyenburg, K., Nielsen, L.D., Johnsen, K., Molin, S., Svenningsen, B., and Miozzari, G. (1978) Reevealuation of the Mode of Action of Streptolydigin in *Escherichia coli*: Induction of Transcription Termination In Vivo. *Antimicrobial Agents and Chemotherapy*, **13**, 234-243.

59. Cassani, G., Burgess, R.R., Goodman, H.M., Gold, L. (1971) Inhibition of RNA polymerase by streptolydigin. *Nat. New. Biol.*, **230**, 197-200.

60. Tuske, S., Sarafianos, S.G., Wang, X., Hudson, B., Sineva, E., Mukhopadhyay, J., Birktoft, J.J., Leroy, O., Ismail, S., Clark, A.D., Dharia, C., Napoli, A., Laptenko, O., Lee, J., Borukhov, S., Ebright, R.H., Arnold, E. (2005) Inhibition of Bacterial RNA Polymerase by Streptolydigin: Stabilization of a straight-bridge-helix active-center conformation. *Cell*, **122**, 541–552.

61. Temiakov, D., Zenkin, N., Vassylyeva, M. N., Perederina, A., Tahirov, T.H., Kashkina, E., Savkina, M., Zorov, S., Nikiforov, V., Igarashi, N., Matsugaki., N., Wakatsuki, S., Severinov, K., Vassylyev, D.G. (2005) Structural basis of transcription inhibition by antibiotic streptolydigin. *Mol. Cell*, **19**, 655-666.

62. Paramanathan, T., Vladescu, I., McCauley, M.J., Rouzina, I., and Williams, M.C. (2012) Force spectroscopy reveals the DNA structural dynamics that govern the slow binding of Actinomycin D. *Nucleic Acids Res.*, **40**, 4925–4932.

63. Sobell, H.M., Jain, S.C., Sakore, T.D. and Nordman, C.E. (1971) Stereochemistry of Actinomycin–DNA Binding. *Nat. New Biol.*, **231**, 200-205.

64. Sobell, H.M. (1985) Actinomycin and DNA transcription. *Proc. Natl. Acad. Sci. U. S. A.*, **82**, 5328-5331.

65. Lian, C., Robinson, H. and Wang, A.H.J. (1996) Structure of Actinomycin D bound with (GAAGCTTC)2 and (GATGCTTC)2 and its binding to the (CAG)n:(CTG)n triplet sequence as determined by NMR analysis. *J. Am. Chem. Soc.*, **118**, 8791-8801.

66. Selby, C.P., Sancar, A. (1993) Molecular mechanism of transcription coupling repair. *Science*, **260**, 53-58.

67. Deaconescu, A.M., Chambers, A.L., Smith, A.J., Nickels, B.E., Hochschild, A., Nigel J. Savery, N.J., Darst, S.A. (2006) Structural basis for bacterial transcription-coupled DNA repair. *Cell*, **124**, 507–520.

68. Roberts, J., Park, J-S. (2004) Mfd, the bacterial transcription repair coupling factor: translocation, repair, and termination. *Current Opinion in Microbiology*, **7**, 120-125.

69. Selby, C.P., Sancar, A. (1995) Structure and function of transcription-repair coupling factor. II. Catalytic properties. *J. Biol. Chem.*, **270**, 4890-4895.

70. Boubakri, H., de Septenville, A.L., Viguera, E., Michel, B. (2010) The helicases DinG, Rep, and UvrD cooperate to promote replication across transcription units in vivo. *EMBO J.*, **29**, 145-157.

71. Guy, C.P., Atkinson, J., Gupta, M.K., Mahdi, A.A., Gwynn, E.J., Rudolph, C.J., Moon, P.B., van Knippenberg, I.C., Cadman, C.J., Dillingham, M.S. *et al.* (2009) Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol. Cell*, **36**, 654-666.

72. Kumari, A., Minko, I.G., Smith, R.L., Lloyd, R.S. and McCullough, A.K. (2010) Modulation of UvrD helicase activity by covalent DNA-protein cross-links. *J. Biol. Chem.*, **285**, 21313-21322.

73. Trautinger, B.W., Jakataji, R.P., Rusakova, E. and Lloyd, R.G. (2005) RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol. Cell*, **19**, 247–258.

74. Washburn, R.S., Gottesman, M.E. (2011) Transcription termination maintains chromosome integrity. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 792–797.

75. Pavco, P.A., Steege, D.A. (1990) Elongation by *Escherichia coli* RNA polymerase is blocked in vitro by a site-specific DNA binding protein. *J. Biol. Chem.*, **265**, 9960-9969.

76. Sukhodolets, M.V., Cabrera, J.E., Zhi, H. and Jin, D.J. (2001) RapA, a bacterial homolog of SWI2/SNF2, stimulates RNA polymerase recycling in transcription. *Genes Dev.*, **15**, 3330-3341.

77. Muzzin, O., Campbell, E.A., Xia, L., Severinova, E., Darst, S.A, and Severinov, K. (1998) Disruption of *Escherichia coli* HepA, an RNA polymerase-associated protein, causes UV sensitivity. *J. Biol. Chem.*, **273**, 15157-15161.

78. Orlova, M., Newlands, J., Das, A., Goldfarb, A., and Borukhov, S. (1995) Intrinsic transcript cleavage activity of RNA polymerase. *Proc. NatI. Acad. Sci. U.S.A.*, **92**, 4596-4600.

79. Borukhov, S., Polyakov, A.,Nikiforov, V., and Goldfarb, A. (1992) GreA protein: A transcription elongation factor from *Escherichia coli*. *Proc. NatI. Acad. Sci. U.S.A.*, **89**, 8899-8902.

80. Toulme, F., Mosrin-Huaman, C., Sparkowski, J., Das, A., Leng, M., and Rahmouni, A. R. (2000) GreA and GreB proteins revive backtracked RNA polymerase in vivo by promoting transcript trimming. *EMBO J.*, **19**, 6853-6859.

81. Tehranchi, A.K., Blankschien, M.D., Zhang, Y., Halliday, J.A., Srivatsan, A., Peng, J., Herman, C., Wang, J.D. (2010) The transcription factor DksA prevents conflicts between DNA replication and transcription machinery. *Cell*, **141**, 595-605.

82. Epshtein, V., Toulme, F., Rahmouni, A.R., Borukhov, S., and Nudler, E. (2003) Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J.*, **22**, 4719-4727.

83. Krasich, R., Wu, S.Y., Kuo, H.K. and Kreuzer, K.N. (Unpublished) Functions that protect Escherichia coli from DNA-protein crosslinks.

84. Kuo, H.K., Krasich, R., Bhagwat, A.S., Kreuzer, K.N. (2010) Importance of the tmRNA system for cell survival when transcription is blocked by DNA-protein cross-links. *Mol. Microbiol.*, **78**, 686-700.

85. Keiler, K.C., Waller, P.R.H., Sauer, R.T. (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*, **271**, 990-993.

86. Roche, E.D. and Sauer, R.T. (2001) Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons. *J. Biol. Chem.*, **276**, 28509–28515.

87. Bandaru, B., Gopal, J. and Bhagwat, A.S. (1996) Overproduction of DNA cytosine methyltransferases causes methylation and C --> T mutations at non-canonical sites. *J. Biol. Chem.*, **271**, 7851-7859.

88. Roche, E.D., Sauer, R.T. (1999) SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. *EMBO J.*, **18**, 4579-4589.

89. Sunohara, T., Abo, T., Inada, T. and Aiba, H. (2002) The C-terminal amino acid sequence of nascent peptide is a major determinant of SsrA tagging at all three stop codons. *RNA*, **8**, 1416–1427.

90. Garza-Sanchez, F., Shoji, S., Fredrick, K., Hayes, C.S. (2009) RNaseII is important for A-site mRNA cleavage during ribosome pausing. *Mol. Microbiol.*, **73**, 882-897.

91. Selby, C.P. and Sancar, A. (1994) Mechanisms of Transcription-Repair Coupling and Mutation Frequency Decline. *Microbiological Reviews*, **58**, 317-329.

92. Zwiefka, A., Kohn, H. and Widger, W.R. (1993) Transcription termination factor Rho: the site of bicyclomycin inhibition in *E. coli*. *Biochemistry*, **32**, 3564–3570.

93. Berenbaum, M.C. (1978) A method for testing for synergy with any number of agents. *J. Infect. Dis.*, **137**, 122–130.

94. Botelho, M.G. (2000) Fractional inhibitory concentration index of combinations of antibacterial agents against cariogenic organisms. *Journal of Dentistry*, **28**, 565-570.

95. Gupta, M.K., Guy, C.P., Yeeles, J.T., Atkinson, J., Bell, H., Lloyd, R.G., Marians, K.J. and McGlynn, P. (2013) Protein--DNA complexes are the primary sources of replication fork pausing in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 7252-7257.

96. Choy, J.S., Aung, L.L.,1 and Karzai, A.W. (2007) Lon Protease Degrades Transfer-Messenger RNA-Tagged Proteins. *J. Bacteriol.*, **189**, 6564–6571.

97. Gottesman, S., Roche, E., Zhou, Y., Sauer, R.T. (1998) The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev*, **12**, 1338-1347.

98. Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.*, **177**, 4121-4130.

99. Jin, J., Bai, L., Johnson, D. S., Fulbright, R.M., Kireeva, M. L., Kashlev, M., Wang, M.D. (2010) Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. *Nat. Struct. Mol. Biol.*, **17**, 745-752.

100. Nakano, T., Ouchi, R., Kawazoe, J., Pack, S.P., Makino, K. and Ide, H. (2012) T7 RNA polymerases backed up by covalently trapped proteins catalyze highly error prone transcription. *J. Biol. Chem.*, **287**, 6562-6572.

101. Kalman, M., Murphy, H., and Cashel, M. (1992) The nucleotide sequence of *recG*, the distal *spo* operon gene in *Escherichia coli* K-12. *Gene*, **110**, 95-99.

102. Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto,E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (A Complete Set of *E. coli* K-12 ORF Archive): Unique resources for biological research. *DNA Research*, **12**, 291–299.

103. Heitman, J., and Model, P. (1991) SOS induction as an in vivo assay of enzyme-DNA interactions. *Gene*, **103**, 1-9.

104. Kenyon, C.J., and Walker, G.C. (1980) DNA-damaging agents stimulate gene expression at specific loci in *Escherichia coli*. *Proc. Natl. Acad. U.S.A.*, **77**, 2819-2823.

105. Barron, C., and Bade, E.G. (1988) Transcriptional mapping of the bacteriophage Mu DNA. *J. Gen. Virol.*, **69**, 385-393.

106. Akroyd, J., and Symonds, N. (1986) Localization of the *gam* gene of bacteriophage mu and characterisation of the gene product. *Gene*, **49**, 173-182.

107. Krause, H.M., Rothwell, M.R., and Higgins, N.P. (1983) The early promoter of bacteriophage Mu: definition of the site of transcript initiation. *Nucleic Acids Res.*, **11**, 5483-5495.

108. Abraham, Z.H.L., Symonds, N. (1990) Purification of overexpressed *gam* gene protein from bacteriophage Mu by denaturation-renaturation techniques and a study of its DNA-binding properties. *Biochem. J.*, **269**, 679-684.

109. Shee, C., Cox, B.D., Gu, F., Luengas, E.M., Joshi, M.C., Chiu, L.Y., Magnan, D., Halliday, J.A., Frisch, R.L., Gibson, J.L. *et al.* (2013) Engineered proteins detect spontaneous DNA breakage in human and bacterial cells. *eLife*, **2**, e01222.

110. Casadaban, M.J., and Cohen, S.N. (1979) Lactose genes fused to exogenous promoters in one step using a Mu-*lac* bacteriophage: *In vivo* probe for transcriptional control sequences. *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 4530-4533.

111. Heitman, J. and Model, P. (1987) Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. *J. Bacteriol.*, **169**, 3243-3250.

112. Hübner, P., Iida, S., Arber, W. (1987) A transcriptional terminator sequence in the prokaryotic transposable element IS1. *Mol. Gen. Genet.*, **206**.

113. Maenhaut-Michel, G., Blake, C.E., Leach, D.R.F., and Shapiro, J.A. (1997) Different structures of selected and unselected araB-lacZ fusions. *Mol. Microbiol.*, **23**, 1133-1145.

114. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.

115. Connelly, J.C., de Leau, E.S., Leach, D.R. (2003) Nucleolytic processing of a protein-bound DNA end by the *E. coli* SbcCD (MR) complex. *DNA Repair*, **2**, 795-807.

116. Neale, M.J., Pan, J. and Keeney, S. (2005) Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature (London)*, **436**, 1053-1057.

117. Schlacher, K., Pham, P., Cox, M.M., Goodman, M.F. (2006) Roles of DNA polymerase V and RecA protein in SOS damage-induced mutation. *Chem. Rev.*, **106**, 406-419.

118. Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A. and Ellenberger, T. (2006) *DNA Repair and Mutagenesis*. Second ed. ASM Press, Washington, DC.

119. Courcelle, J., Carswell-Crumpton, C. and Hanawalt, P.C. (1997) recF and recR are required for the resumption of replication at DNA replication forks in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A*, **94**, 3714-3719.

120. Courcelle, J., Donaldson, J.R., Chow, K.H. and Courcelle, C.T. (2003) DNA damage-induced replication fork regression and processing in *Escherichia coli*. *Science*, **299**, 1064-1067.

121. Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K., Hartl, F.U. (1992) Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature*, **356**, 683-689.

122. Schröder, H., Langer, T., Hartl, F.U., and Bukau, B. (1993) DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBO J.*, **12**.

123. Hayes, C.S., Holberger, L.E. (2009) Ribosomal Protein S12 and Aminoglycoside Antibiotics Modulate A-site mRNA Cleavage and Transfer-Messenger RNA Activity in *Escherichia coli*. *J. Biol. Chem.*, **284**, 32188-32200.

124. Hayes, C.S., and Sauer, R.T. (2003) Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. *Mol. Cell*, **12**, 903-911.

125. Chambers, A.L., Smith, A.J. and Savery, N.J. (2003) A DNA translocation motif in the bacterial transcription – repair coupling factor, Mfd. *Nucleic Acids Res.*, **31**, 6409–6418.

126. Jin, D.J., Zhou, Y.N., Shaw, G. and Ji, X. (2011) Structure and function of RapA: a bacterial Swi2/Snf2 protein required for RNA polymerase recycling in transcription. *Biochim. Biophys. Acta.*, **1809**, 470-475.

127. Nakano, T., Ouchi, R., Kawazoe, J., Pack, S.P., Makino, K., Ide, H. (2012) T7 RNA polymerases backed up by covalently trapped proteins catalyze highly error prone transcription. *J. Biol. Chem.*, **287**, 6562-6572.

128. Fish, R.N. and Kane, C.M. (2002) Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta.*, **1577**, 287-307.

129. Seigneur, M., Bidnenko, V., Ehrlich, S.D. and Michel, B. (1998) RuvAB acts at arrested replication forks. *Cell*, **95**, 419-430.

130. Hong, G. and Kreuzer, K.N. (2003) Endonuclease cleavage of blocked replication forks: An indirect pathway of DNA damage from antitumor drug-topoisomerase complexes. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 5046-5051.

131. Connelly, J.C., Kirkham, L.A. and Leach, D.R.F. (1998) The SbcCD nuclease of *Escherichia coli* is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 7969-7974.

132. Connelly, J.C. and Leach, D.R. (2002) Tethering on the brink: the evolutionarily conserved Mre11-Rad50 complex. *Trends in Biochemical Sciences*, **27**, 410-418.

133. Straus, D., Walter, W., and Gross, C.A. (1990) DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of sigma 32. *Genes Dev.*, **4**, 2202-2209.

134. Whitney, E. (1971) The tolC locus in *Escherichia coli* K12. *Genetics*, **67**, 39–53.

135. Leivel, L. (1965) Actinomycin sensitivity of *Escherichia coli* produced by EDTA. *Biochem. Biophys. Res. Commun.*, **28**, 229-236.

136. Erie, D.A., Hajiseyedjavadi, O., Young, M.C. and von Hippel, P.H. (1993) Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science*, **262**, 867-873

137. Fisher, R.F. and Yanofsky, C. (1983) Mutations of the b subunit of RNA polymerase alter both transcription pausing and transcription termination in the trp operon leader region *in vitro*. *J. Biol. Chem.*, **258**, 8146-8150.

138. Jin, D.J., Walter, W. A., Gross, C. A. (1988) Characterization of the termination phenotypes of rifampicin-resistant mutants. *J. Mol. Biol.*, **202**, 245–253.

139. Neff, N.F., and Chamberlin, M.J. (1980) Termination of transcription by Escherichia coli ribonucleic acid polymerase in vitro. Effect of altered reaction conditions and mutations in the enzyme protein on termination with T7 and T3 deoxyribonucleic acids. *Biochemistry*, **19**, 3005-3015.

140. Jun Jin, D., and Gross, C.A. (1988) Mapping and Sequencing of Mutations in the Escherichia coli rpoB Gene that Lead to Rifampicin Resistance. *J. Mol. Biol.*, **202**, 45-58.

141. Mirkin, E.V. and Mirkin, S.M. (2005) Mechanisms of transcription-replication collisions in bacteria. *Mol. Cell. Biol.*, **25**, 888–895.

142. Freudenreich, C.H. and Kreuzer, K.N. (1993) Mutational analysis of a type II topoisomerase cleavage site: distinct requirements for enzyme and inhibitors. *EMBO J.*, **12**, 2085-2097.

143. Freudenreich, C.H. and Kreuzer, K.N. (1994) Localization of an aminoacridine antitumor agent in a type II topoisomerase-DNA complex. *Proc Natl Acad Sci U S A.*, **91**, 11007-11011.

144. Freudenreich, C.H., Chang, C. and Kreuzer, K.N. (1998) Mutations of the bacteriophage T4 type II DNA topoisomerase that alter sensitivity to antitumor agent 4'-(9-acridinylamino)methanesulfon-m-anisidide and an antibacterial quinolone. *Cancer Res.*, **58**, 1260-1267.

145. Dabholkar, M., Thornton, K., Vionnet, J., Bostick-Bruton, F., Yu, J.J. and Reed, E. (2000) Increased mRNA levels of xeroderma pigmentosum complementation group B (XPB) and Cockayne's Syndrome complementation group B (CSB) without increased mRNA levels of multidrug-resistance gene (MDR1) or metallothionein-II (MT-II) in platinum-resistant human ovarian cancer tissues. *Biochemical Pharmacology*, **60**, 1611–1619.

146. Baharoglu, Z., Babosan, A. and Mazel, D. (2013) Identification of genes involved in low aminoglycoside-induced SOS response in *Vibrio cholerae*: a role for transcription stalling and Mfd helicase. *Nucleic Acids Res.*, **[published ahead of print]**.

147. Park, J., Marr, M.T., Roberts, J.W. (2002) *E.coli* transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell*, **109**, 757-767.

148. Woudstra, E.C., Gilbert, C., Fellows, J., Jansen, L., Brouwer, J., Erdjument-Bromage[3], H., Tempst, P. and Svejstrup, J.Q. (2002) A Rad26±Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature*, **415**, 929-933.

149. McKaya, B.C., Chenb, F., Clarkeb, S.T., Wigginb, H.E., Harleyb, L.M. and Ljungman, M. (2001) UV light-induced degradation of RNA polymerase II is dependent on the Cockayne's syndrome A and B proteins but not p53 or MLH1. *Mutat. Res.*, **485**, 93–105.

150. Krugh, T.R. (1972) Association of actinomycin D and deoxyribodinucleotides as a model for binding of the drug to DNA. *Proc. Natl Acad. Sci. U.S.A.*, **69**.

151. George, D.L. and Witkin, E.M. (1974) Slow excision repair in an *mfd* mutant of *Escherichia coli* B/r. *Mol Gen Genet.* , **133**, 283--291.

152. George, D.L. and Witkin, E.M. (1975) Ultraviolet light-induced responses of an mfd mutant of Escherichia coli B/r having a slow rate of dimer excision. *Mutat. Res.*, **28**, 347-354.

153. Hanawalt, P., C. (2002) Subpathways of nucleotide excision repair and their regulation. *Oncogene*, **21**, 8949 – 8956.

154. Chan, C.L., Wang, D. and Landick, R. (1997) Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. *J. Mol. Biol.*, **268**, 54-68.

155. Artsimovitch, I. and Landick, R. (1998) Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. *Genes Dev.*, **12**, 3110-3122.

156. Guajardo, R. and Sousa, R. (1997) A model for the mechanism of polymerase translocation. *J. Mol. Biol.*, **265**, 8-19.

157. Komissarova, N. and Kashlev, M. (1997) RNA polymerase switches between inactivated and activated states By translocating back and forth along the DNA and the RNA. *J. Biol. Chem.*, **272**, 15329-15338.

158. Nudler, E., Mustaev, A., Lukhtanov, E. and Goldfarb, A. (1997) The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell*, **89**, 33–41.

159. Branum, M.E., Reardon, J.T. and Sancar A. . (2001) DNA repair excision nuclease attacks undamaged DNA. A potential source of spontaneous mutations. *J. Biol. Chem.*, **276**, 25421-25426.

160. Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao,Y. (1997) The complete genome sequence of Escherichia coli K-12. *Science*, **277**, 1453-1474.

161. Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc.*, **4**, 44-57.

162. Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1-13.

163. Salgado, H., Peralta-Gil, M., Gama-Castro, S., Santos-Zavaleta, A., Muñiz-Rascado, L., García-Sotelo, J.S., Weiss, V., Solano-Lira, H., Martínez-Flores, I., Medina-Rivera, A. *et al.* (2012) RegulonDB (version 8.0): Omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res*, **41**, D203–D213.

164. Keseler, I.M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., Fulcher, C., Huerta, A.M., Kothari, A., Krummenacker, M. *et al.* (2013) EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res*, **41**, D605-612.

165. Nomura, M., Gourse, R. and Baughman, G. (1984) Regulation of the synthesis of ribosomes and ribosomal components. *Annu Rev Bichem*, **53**, 75–117.

166. Bremer, H. and Dennis, P. (1996) In Neidhardt, F. (ed.), *Escherichia coli and Salmonella*. ASM Press, Washington, DC:, Vol. 2, pp. 1553-1569.

167. Keener, J.N., M. (1996) In Neidhardt, F. (ed.), *Escherichia coli and Salmonella*. ASM Press, Washington, DC, Vol. 1, pp. 1417-1431.

168. Gausing, K. (1977) Regulation of ribosome production in Escherichia coli: synthesis and stability of ribosomal RNA and of ribosomal protein messenger RNA at different growth rates. *J. Mol. Biol.*, **115**, 335–354.

169. Jin, D.J., Cagliero, C. and Zhou, Y.N. (2012) Growth rate regulation in Escherichia coli. *FEMS Microbiol Rev*, **36**, 269–287.

170. Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W. and Gourse, R.L. (2004) DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell*, **118**, 311–322.

171. Aberg, A., Fernandez-Vazquez, J., Cabrer-Panes, J.D., Sanchez, A. and Balsalobre, C. (2009) Similar and divergent effects of ppGpp and dksA deficiencies on transcription in *Escherichia coli*. *J. Bacteriol.*, **191**, 3226-3236.

172. Haddadin, F.T., and Harcum, S.W. (2005) Transcriptome Profiles For High-Cell-Density Recombinant and Wild-Type Escherichia coli. *Biotechnol Bioeng.*, **90**, 127-153.

173. Bradley, M.D., Beach, M.B., de Koning, A.P., Pratt, T.S. and Osuna, R. (2007) Effects of Fis on Escherichia coli gene expression during different growth stages. *Microbiology*, **153**, 2922-2940.

174. Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J.P., Danchin, A. and Bertin, P. (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol. Microbiol.*, **40**, 20-36.

175. Cardinale, C.J., Washburn, R.S., Tadigotla, V.R., Brown, L.M., Gottesman, M.E. and Nudler, E. (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science*, **320**, 935-938.

176. Menouni, R., Champ, S., Espinosa, L., Boudvillain, M. and Ansaldi, M. (2013) Transcription termination controls prophage maintenance in Escherichia coli genomes. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 14414-14419. .

177. Zeng, X., Galinier, A. and Saxild, H.H. (2000) Catabolite repression of dra–nupC–pdp operon expression in Bacillus subtilis. *Microbiology*, **146**, 2901–2908.

178. Grossman, A.D., Straus, D.B., Walter, W.A. and Gross, C.A. (1987) Sigma 32 synthesis can regulate the synthesis of heat shock proteins in Escherichia coli. *Genes Dev.*, **1**, 179-184.

179. Nishino, K. and Yamaguchi, A. (2001) Analysis of a complete library of putative drug transporter genes in *Escherichia coli*. *J. Bacteriol.*, **183**, 5803-5812.

180. Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L. and Keasling, J.D. (2001) Homogeneous expression of the P(BAD) promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology*, **147**, 3241-3247.

Biography

Rachel Krasich was born on March 8, 1986 in St. Louis Park, Minnesota to John and Mary Krasich. She graduated from Westfield High School in 2004 and attended college at Rose-Hulman Institute of Technology. In 2008, she graduated *summa cum laude* with a Bachelor of Science in Biochemistry and Molecular Biology, a Bachelor of Science in Chemistry, and a Minor in Economics. During her undergraduate years she conducted research at the Wells Center for Pediatric Research with Dr. David Ingram on endothelial colony forming cells. Following graduation, Rachel attended Duke University to pursue of her doctoral degree in Dr. Kenneth Kreuzer's laboratory. While at Duke, her work has focused on transcriptional elongation inhibition.

Publications:

Kuo, H. K., **Krasich, R**., Bhagwat, A.S., Kreuzer, K.N. (2010) Importance of the tmRNA system for cell survival when transcription is blocked by DNA-protein crosslinks, *Molecular Microbiology 78*(3):686-700.

Yoder, M.C., Mead, L.E., Prater, D., Krier, T.R., Mroueh, K.N., Li, F**., Krasich, R**., Temm, C.J., Prchal, J.T., Ingram, D.A. (2007) Redefining endothelial progenitor cells via clonal analysis and hematopoietic stem/progenitor cell principals, *Blood* 109(5):1801-9.

Manuscripts in Progress:

Krasich, R., Wu, Y., Kuo, H.K., Kreuzer, K.N. (In Submission) Functions that protect *Escherichia coli* from DNA-protein crosslinks

Krasich, R., Al-Khalil, R**.**, Kreuzer, K.N. (In progress) Fate of transcription elongation complexes stalled by exogenous elongation inhibitors

Krasich, R, Kreuzer, K.N. (In progress) Novel role of Mfd in undamaged cells